

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : A01H 5/00, C07H 21/04, C12N 5/14, 15/29, 15/52, 15/82		A1	(11) International Publication Number: WO 00/67558 (43) International Publication Date: 16 November 2000 (16.11.00)
(21) International Application Number: PCT/US00/12450 (22) International Filing Date: 5 May 2000 (05.05.00) (30) Priority Data: 60/132,919 6 May 1999 (06.05.99) US		(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(71)(72) Applicant and Inventor: TIMKO, Michael [US/US]; 1610 Old Ballard Road, Charlottesville, VA 22901 (US). (74) Agent: HANSEN, Christine, M.; Connolly Bove Lodge & Hutz LLP, 1210 Market Street, P.O. Box 2207, Wilmington, DE 19899 (US).		Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
<p>(54) Title: REGULATION OF GENE EXPRESSION IN TOBACCO FOR MANIPULATION OF PLANT GROWTH AND SECONDARY METABOLISM</p> <p>(57) Abstract</p> <p>This invention relates to enzymes involved in alkaloid, and specifically nicotine, formation in tobacco plants. The invention is based, at least in part, on the nucleotide sequences encoding four variants of putrescine N-methyltransferase (PMT1, PMT2, PMT3, and PMT4), two variants of arginine decarboxylase (ADC1 and ADC2), ornithine decarboxylase (ODC), S-adenosylmethionine synthetase (SAMS), a fragment of NADH dehydrogenase, and a fragment of phosphoribosylanthranilate isomerase. The invention also relates to proteins expressed by these nucleotides, promoter regions of these nucleotides, use of these promoter regions to culture transgenic plant cells and to produce transgenic plants, sense and antisense nucleotides complementary to all or portions of these nucleotide sequences, use of sense and antisense nucleotides to regulate gene expression, and assays using proteins involved in alkaloid formation in tobacco plants.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Larvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LJ	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**REGULATION OF GENE EXPRESSION IN TOBACCO FOR MANIPULATION OF
PLANT GROWTH AND SECONDARY METABOLISM**

CROSS-REFERENCE TO RELATED APPLICATIONS

5 This application is a continuation-in-part of US Patent Application Ser. No. 60/ 132,919, filed May 6, 1999, now abandoned, which is hereby incorporated by reference in its entirety herein.

FIELD OF THE INVENTION

This invention relates to enzymes involved in alkaloid, and specifically nicotine, formation in 10 tobacco plants. The invention is based, at least in part, on the nucleotide sequences encoding four variants of putrescine N-methyltransferase (PMT1, PMT2, PMT3, and PMT4), two variants of arginine decarboxylase (ADC 1 and ADC2), ornithine décarboxylase (ODC), S-adenosylmethionine synthetase (SAMS), a fragment of NADH dehydrogenase, and a fragment of phosphoribosylanthranilate isomerase. The invention also relates to proteins expressed by these 15 nucleotides, promoter regions of these nucleotides, use of these promoter regions to culture transgenic plant cells and to produce transgenic plants, sense and antisense nucleotides complementary to all or portions of these nucleotide sequences, use of sense and antisense nucleotides to regulate gene expression, and assays using proteins involved in alkaloid formation in tobacco plants.

20

BACKGROUND OF THE INVENTION

I. Alkaloid Formation

Alkaloids are one of the most diverse groups of secondary compounds found in plants and they are the product of a complex biosynthesis pathway (Hashimoto and Yamada, 1994; Chou and 25 Kutchan, 1998; Waterman, 1998). Why plants accumulate these compounds and in so many different forms is not known. Moreover, for many alkaloids, the exact site of synthesis and the factors that control their intercellular distribution and accumulation remain to be determined (Hashimoto and Yamada, 1994; Kutchan, 1995; Chou and Kutchan, 1998).

Nicotine is the most abundant alkaloid present in cultivated tobacco. Nicotine is formed 30 primarily in the roots of the tobacco plant and subsequently is transported to the leaves, where it is stored (Tso, Physiology and Biochemistry of Tobacco Plants, pp. 233-34, Dowden, Hutchinson & Ross, Stroudsburg, Pa. (1972)).

The synthesis and accumulation of nicotine and other tobacco alkaloids are known to be controlled by various developmental, environmental, and chemical cues. Changes in phytohormone

(e.g., auxin, cytokinin) levels and/or ratios as a consequence of developmental age (Hashimoto and Yamada, 1994; Kutchan, 1995) or by direct manipulation of plant cell culture conditions have been shown to affect the synthesis and accumulation of nicotine and various tobacco alkaloids (Hashimoto and Yamada, 1994; Hibi *et al.*, 1994; Eilbert, 1998). Various abiotic factors (wounding, drought stress, pH imbalance, etc.) [Hashimoto and Yamada, 1994; Kutchan, 1998; Waterman, 1998] 1, 2, 4], as well as biotic factors, such as herbivory, insect feeding, and attack by various microbial and fungal pathogens, are known elicit increased production of nicotine and other alkaloids in the leaves of wild and cultivated tobacco species (Baldwin, 1989; Saito and Murakoishi, 1998; Baldwin and Prestin, 1999). In addition, the commercial practice of topping (i.e., removal of flowering head and young leaves at the upper portions of the plant), results in increases in nicotine and the amount and complexity total alkaloids present in the leaves of *Nicotiana tabacum* (Hashimoto and Yamada, 1994; Hibi *et al.*, 1994). The factors controlling the topping-induced increase in alkaloid biosynthesis are not known, but likely involve a complex physiological response in the plant as a result of altered phytohormones and wound induced signaling (Akehurst, 1981; Hibi *et al.*, 1994; Kutchan, 1998). In this regard, considerable evidence now exists indicating that a jasmonic acid (JA)-mediated signal transduction pathway may play a role in regulation of gene expression contributing to this increase in alkaloid biosynthesis (Baldwin *et al.*, 1994, 1996, 1997; Ohnmeiss *et al.*, 1997; Imanishi *et al.*, 1998a, 1998b).

The nicotine molecule is comprised of two heterocyclic rings, a pyridine moiety and a 20 pyrrolidine moiety, each of which is derived from a separate biochemical pathway. The pyridine moiety of nicotine is derived from nicotinic acid. The pyrrolidine moiety of nicotine is provided through a pathway leading from putrescine to N-methylputrescine and then to N-methylpyrrolidine. (Goodwin and Mercer, Introduction to Plant Biochemistry, pp. 488-91, Pergamon Press, New York, (1983)).

Putrescine is formed in plants by one of two pathways (Chattopadhyay and Ghosh, 1998). It can be synthesized directly from ornithine, in a reaction catalyzed by the enzyme ornithine decarboxylase (ODC, EC 4.1.1.17), or formed indirectly from arginine in a reaction sequence initiated by arginine decarboxylase (ADC, EC 4.1.1.19). Putrescine formed by the ADC and/or ODC pathway serves as precursor in the synthesis of the higher polyamines, spermine and spermidine, 30 catalyzed by the enzymes spermine synthase and spermidine synthase, respectively, or it is converted to N-methylputrescine by the action of putrescine N-methyltransferase (PMT), the first committed step in nicotine biosynthesis (Hashimoto and Yamada, 1994; Kutchan, 1995; Chattopadhyay and Ghosh, 1998). N-methyl putrescine is oxidized by a diamine oxidase and cyclized to form the 1-methyl- Δ^1 -pyrrolium cation, which is condensed with nicotinic acid or its derivative to form nicotine

(Hashimoto and Yamada, 1994).

Putrescine is a precursor for N-methylputrescine, which then forms N-methylpyrrolidine. Conversion of putrescine to N-methylputrescine is catalyzed by the enzyme putrescine N-methyltransferase ("PMT"), with S-adenosylmethionine serving as the methyl group donor. PMT appears to be the rate-limiting enzyme in the pathway supplying N-methylpyrrolidine for nicotine synthesis in tobacco (Feth et al., "Regulation in Tobacco Callus of Enzyme Activities of the Nicotine Pathway", *Planta*, 168, pp. 402-07 (1986); Wagner et al., "The Regulation of Enzyme Activities of the Nicotine Pathway in Tobacco", *Physiol. Plant.*, 68, pp. 667-72 (1986)).

10 II. TRANSGENIC PLANTS

The methods of nicotine formation in tobacco and the genes involved have been studied both to better understand differential gene expression during tobacco growth and development, and also to discover tools useful for creating transgenic plants. For example, the regulatory sequences that modify protein expression in tobacco may be useful in creating transgenic tobacco or other
15 transgenic plants.

It has already been demonstrated that tissues of many plant species may be transformed by exogenous, typically chimeric, genes which are effective to stably transform cells of the tissues. For several species, tissues transformed in this fashion may be regenerated to give rise to whole transgenic or genetically engineered plants. The engineered traits introduced into the transgenic plants by these techniques have proven to be stable and have also proven to be transmissible through normal Mendelian inheritance to the progeny of the regenerated plants. One such desirable trait is the production in the plant cells of desired gene products *in vivo* in the cells of the transgenic plants. For a chimeric gene to be effective, the foreign DNA sequence containing a coding region should be flanked by appropriate promotion and control regions. Commonly used plant cell transcription
20 promoters include the nopaline synthase promoter from the T-DNA of *A. tumefaciens* and the 35S promoter from the cauliflower mosaic virus.
25

In order for the newly inserted chimeric gene to express the protein for which it codes in the plant cell, the proper regulatory signals must be present and in the proper location with respect to the gene. These regulatory signals include a promoter region, a 5' non-translated leader sequence and a 3' polyadenylation sequence. A promoter is a DNA sequence that directs the cellular machinery of a plant to produce RNA from the contiguous structural coding sequence downstream (3') to the promoter. The promoter region influences the rate at which the RNA product of the gene and resultant protein product of the gene is made. The 3' polyadenylation signal is a non-translated region that functions in

the plant cells to cause the addition of polyadenylate nucleotides to the 3' end of the RNA to enable the mRNA to be transported to the cytoplasm and to stabilize the mRNA for subsequent translation of the RNA to produce protein.

Other plant cell transformation techniques are directed toward the direct insertion of DNA into the cytoplasm of plant cells from which it is taken up, by an uncharacterized mechanism, into the genome of the plant. One such technique is electroporation, in which electric shock causes disruption of the cellular membranes of individual plant cells. Plant protoplasts in aqueous solution when subject to electroporation will uptake DNA from the surrounding medium. Another technique involves the physical acceleration of DNA, coated onto small inert particles, either into regenerable plant tissues or into plant germline cells.

The availability of cloned nucleic acid sequences encoding an enzyme involved in alkaloid synthesis allows for the potential manipulation of alkaloid contents. Furthermore, the availability of promoters useful for expressing genes in plants allows for the creation of chimeric molecules and transgenic plants, which in turn result in possible native plant production of desirable proteins.

Previously reported work discloses cloning nucleotides encoding proteins involved in the biosynthesis of nicotine, and isolating such proteins. Approximately twenty or more cDNAs and/or genomic DNA fragments encoding different enzymes involved with alkaloid formation have been isolated (Chattopadhyay and Ghosh, 1998). For example, successful cloning of partial or full-length cDNA encoding ODC activity from tobacco was disclosed by (Malik *et al.*, *J. Plant Biochem. & Biotech.* 5:109-112 (1996)). Also, a relatively crude preparation of PMT (30-fold purification) has been subjected to limited characterization (Mizusaki *et al.*, "Phytochemical Studies on Tobacco Alkaloids XIV. The Occurrence and Properties of Putrescine N-Methyltransferase in Tobacco Plants", *Plant Cell Physiol.*, 12, pp. 633-40 (1971)). A process for purifying PMT is disclosed in US Patent No. 5,369,023, "Method of purifying putrescine n-methyltransferase from tobacco plant extract with an anion exchange medium", hereby incorporated by reference in its entirety herein. Several laboratories have reported the cloning of partial or full-length cDNAs encoding ADC (Bell and Malmberg, 1990; Rostogi *et al.*, 1993; Perez-Amador *et al.*, 1995; Nam *et al.*, 1997; Watson and Malmberg, 1996). Comparisons of the amino acid sequences of ADC from various plants revealed a high degree of conservation among the various proteins, as well as homology to ODC (Malmberg *et al.*, 1998).

It is an object of the present invention to characterize the nucleotide and amino acid sequences of enzymes involved in the biosynthesis of nicotine in tobacco. It is also an object of the present invention to provide plant promoter regions that are capable of conferring high levels of transcription in rapidly dividing cells of transformed plants when coupled with a heterologous coding

sequence in a chimeric gene. Further, the invention is directed to chimeric genes incorporating such promoter regions, stable transfection of plants with these chimeric genes, and the plants and cells that are transfected, as well as seeds of such transfected plants. It is a further object to characterize sense and antisense nucleotides capable of regulating expression of genes encoding for enzymes involved
5 in the biosynthesis of alkaloids.

SUMMARY OF THE INVENTION

Proteins involved in the biosynthesis of nicotine in tobacco *N. tabacum* are the subject of this invention. More specifically, the invention concerns four variants of putrescine N-methyltransferase
10 (PMT1, PMT2, PMT3, and PMT4), two variants of arginine decarboxylase (ADC 1 and ADC2), ornithine decarboxylase (ODC), S-adenosylmethionine synthetase (SAMS), NADH dehydrogenase, and phosphoribosylanthranilate isomerase.
15

BRIEF DESCRIPTION OF THE FIGURES

Figure 1. Genomic DNA gel blot analysis of the PMT gene family in *N. tabacum* cv. Xanthi.
15 Total genomic DNA (30 µg) was digested with *Kpn*I, *Eco*RI, or *Eco*RI and *Kpn*I, separated by agarose gel electrophoresis, and transferred to nylon membranes. The membrane was hybridized with a ³²P-labeled antisense strand probe covering the complete coding region of the *NtPMT1a* cDNA. Identity of the hybridizing bands as determined by comparison to phage DNA digests is indicated. Molecular weights are given in kb. Note that *Kpn*I shifts only the *NtPMT1b* band in the
20 gel blot because this restriction site is present only in Exon 1 of *NtPMT1b* and not *NtPMT1a*.

Figure 2. Amino acid sequence alignment of *N. tabacum* PMTs. Shown is a PILEUP alignment of
25 the predicted amino acid sequences of the various tobacco PMTs. Amino acid residues that differing among the PMTs are shaded. *NtPMT1a*, *NtPMT2*, *NtPMT3*, and *NtPMT4* refer to the deduced amino acid sequences of the PMTs encoded by the *NtPMT1a*, *NtPMT2*, *NtPMT3*, and *NtPMT4* genes, respectively, isolated from *N. tabacum* cv. Xanthi genomic DNA; c*NtPMT1a* is the predicted amino acid sequence of the A411 cDNA (Accession No. D28506) isolated from *N. tabacum* cv. Burley 21 by Hibi *et al.* (1994). The location of the exon-intron boundaries are indicated by the dark vertical line. The nucleotide sequences for *NtPMT1a*, *NtPMT2*, *NtPMT3*, and *NtPMT4* appear in
30 GenBank under the accession numbers AF126810, AF126809, AF126811, and AF126812, respectively

Figure 3. Polyacrylamide gel electrophoresis analysis of PCR amplified genomic DNA fragments

encoding Exon 1 of PMT from various species of *Nicotiana*. PCR amplification was carried out as described in the Materials and Methods using Exon 1-specific primers 1 and 2 and total genomic DNA isolated from *N. tabacum*, *N. otophora*, and *N. tomentosiformis*. The amplification products were separated by electrophoresis on 6.5% polyacrylamide gels, the gels fixed, and subject to autoradiography. The amplification products isolated from *N. tabacum* cv. Burley 21 and *N. tabacum* cv. Xanthi were identical and only the amplification products from the reactions with *N. tabacum* cv. Burley 21 DNA are shown. Standards were generated in identical reaction conditions primed with plasmid DNA encoding the various *PMT* genes isolated in this study.

10

Figure 4. Nucleotide sequence alignment of the 5'-flanking regions of the *N. tabacum* *PMT* genes. Shown is a PILEUP alignment of the nucleotide sequences upstream of the initiating methionine (MET) codon of the four *PMT* genes isolated from *N. tabacum* cv. Xanthi. The proposed start site for transcription of the *NtPMT1a* gene is indicated by the +1 above the sequences. The TATA-box and CCAAT-box motifs are boxed. Potential transcriptional regulatory elements identified by MOTIF search programs are also boxed and indicated by the following abbreviations: PAL: palindromic sequences; G-Box: G-Box homologous sequences; MRE: metal-responsive element homolog. Nucleotides identical in three or more sequences are shaded. The polyguanine-rich region is underlined. Numbering is indicated to the right and is relative to the proposed start site of each gene.

Figure 5. RNA gel blot analysis of *PMT* transcript levels in various tissues. Total RNA was isolated from various tissues of mature *N. tabacum* cv. Burley 21 and analyzed by gel blot analysis using a ³²P-labeled *NtPMT1a* cDNA coding region (Exons 2 to 8) probe capable of detecting all *PMT* transcripts.

25 A. *PMT* transcript levels in various tobacco plant tissues and/or organs.
B. Induction of *PMT* expression in tobacco roots following topping. Abbreviations: HP, wild-type (*Nic1*/*Nic2*) Burley 21; LP, low alkaloid (*nic1*/*nic2*) mutant. The β-subunit of mitochondrial ATPase (β-ATPase) served as a control.

30 *Figure 6.* Semi-quantitative RT-PCR analysis of *PMT* gene expression in roots of tobacco plant before and after topping.

A. Shown is relative abundance of the individual *PMT* gene transcripts before and after topping. RT-PCR was carried out as described in the Material and methods using Exon 1 specific primers. Messenger RNA was amplified from total RNA isolated from the roots of wild-type (HP,

Nic1Nic2) Burley 21 and low alkaloid (LP, *nic1nic2*) Burley 21 tobacco plants. Far right lane represents size standards for the genes isolated by PCR amplification from plasmid DNA. The β -subunit of mitochondrial ATPase (β -ATPase) served as a control.

5 B. Bar graphs showing relative expression of the individual PMT genes following topping in both HP and LP tobacco roots. Abbreviations: HP, wild-type (*Nic1Nic2*) Burley 21; LP, low alkaloid (*nic1nic2*) mutant.

Figure 7. The nucleotide and predicted amino acid sequences of the transcribed portions of the *N. tabacum* cv Xanthi NtADC1 and NtADC2 genes. Shown are the complete nucleotide and predicted amino acid sequence of the *N. tabacum* cv Xanthi NtADC1 gene and where it differs from the NtADC2 gene sequence. The dots indicate nucleotide sequence identity and the stars indicate amino acid sequence identity. The proposed polyadenylation signal is underlined. The sequences terminate at the point of polyadenylation found in the full length cDNA (Wang, 1999; AF127239). The 10 complete nucleotide sequences for the *N. tabacum* cv Xanthi NtADC1 (AF127240) and NtADC2 15 (AF127241) including the 5' and 3' flanking sequences appear in Genbank.

Fig. 8. Comparison of the predicted amino acid sequences of arginine decarboxylases (ADCs) from various species. Shown is a PILEUP alignment of the predicted amino acid sequence of the *N. tabacum* cv Xanthi NtADC1 gene (AF127240) aligned to the predicted ADC protein sequences from *N. sylvestris* (AB12873), *Arabidopsis thaliana* (AF009647), *Avena sativa* (oat) (X56802), *Lycopersicon esculentum* (tomato) (L16582) and *Escherichia coli* (M31770). Amino acid residues conserved among the various ADC are shaded.

25 Fig. 9. Gel blot analysis of *ADC* transcript levels in the roots of wild-type and low alkaloid mutant Burley 21 tobacco before and after topping. Total RNA was isolated from the roots of mature wild-type and low alkaloid mutant *N. tabacum* cv. Burley 21 and analyzed by gel blot analysis using [α -³²P]-dCTP labeled probes recognizing the coding region of ADC or the β -subunit of tobacco 30 mitochondrial ATP synthase (Boutry and Chua, 1985). Quantitation was carried out by phosphorimaging using a Molecular Dynamics PhosphorImager. Values were normalized relative to the intensities of the *atp2* control band in each lane. The experiment was conducted twice with different total RNA samples.

Fig. 10. Differential expression of NtADC-1 and NtADC-2 in various tissues of tobacco. Expression of the NtADC-1 and NtADC-2 genes was analyzed using semi-quantitative RT-PCR and gene specific primers capable of discriminating between transcripts arising from the two genes. Panel A shows control reactions demonstrating primer specificity in the PCR reactions using plasmids containing the NtADC-1 and NtADC-2 coding sequences. The numbers above the lane refer to the specific primer combinations as described in the Materials and methods. Panel B shows the results of RT-PCR reactions using first strand cDNA synthesized from total RNA extracted from either root, leaf, or flowers. As an internal control, primers specific for the *atp2* gene transcript were included in the amplification reactions. All reactions were carried out within the linear range of template amplification as determined by varying template amount, amplification time, and temperature as described in Riechers and Timko (1999).

Fig. 11. Genomic DNA gel blot analysis of the ODC gene family in *N. tabacum*. Total genomic DNA (30 µg) was digested with *Eco*RI or *Hind*III, fractionated by agarose gel electrophoresis, transferred to nylon membranes and hybridized with an α -³²P-dCTP labeled probe encoding full-length ODC cDNA as described in the Materials. The mobility of molecular weight standards are given to the right of the figure in kilobases (kb).

Fig 12. Comparison of the nucleotide and predicted amino acid sequences of the *NtODC-1* and *NtODC-2* genes. Shown are the nucleotide and predicted amino acid sequences of the *NtODC-1* (AF233850) and *NtODC-2* (AF233849) genes. In the figure, the complete amino acid sequence of the pODC2 is given and the pODC1 sequence is given only where it differs. The start site of transcription is designated as +1 and the poly(A) addition site is indicated by the arrow. Within the relevant regions of homology, nucleotide differences between the *NtODC-1* and *NtODC-2* genes are in bold lettering. The proposed TATA-box, and polyadenylation signal are shaded.

Fig. 13. Protein sequences alignment of ornithine decarboxylases (ODCs) from various species. Shown is a PILEUP alignment of the predicted amino acid sequences of the *N. tabacum* cv. Xanthi pODC2 protein (AF233849) with the ODCs from *N. tabacum* cv. SC58 (Y10472) and cv. BY-2 (ABO31066), *Lycopersicon esculentum* (tomato) (AF030292), *Datura stramonium* (jimsonweed) (X87847), *Saccharomyces cerevisiae* (NP_012737), and humans (*Homo sapiens*; AAA59966). Amino acid residues conserved among the various ODCs are shaded.

Fig. 14. Gel blot analysis of *ODC* transcript levels in various tissues of mature tobacco plants and in the roots before and after topping. Total RNA was isolated from various tissues of mature *N. tabacum* cv. Burley 21 and analyzed by gel blot analysis using an α -³²P-dCTP labeled coding region probes for ODC. (A) Transcript levels in various organs of wild-type tobacco: R, root; S, stem ; L, leaf ; SE, sepal; PE, petal; O, ovary; S, stamen; and AN, anther. (B) Transcript levels in roots of Burley 21 tobacco plants before and after topping. RNA gel blot analysis of the tissues-specific distribution and post-topping expression of transcripts encoding ODC in tobacco. As a control, the blots were also probed with radioactively labeled probes encoding the alkaloid biosynthesis enzyme putrescine N-methyltransferase (PMT) and a root specific β -glucosidase (TBG-1).

10

DETAILED DESCRIPTION OF THE INVENTION

Nucleic acid sequences have been isolated from tobacco that encode important enzymes in nicotine and total alkaloid formation, including PMT1, PMT2, PMT3, PMT4, ADC1, ADC2, ODC, and SAMS. Also identified are cDNA fragments encoding partial segments of NADH dehydrogenase and phosphoribosilanthronilate isomerase. Also identified are promoter regions for the nucleotides encoding PMT1, PMT2, PMT3, PMT4, and ADC2. All of these nucleic acids are isolated from *Nicotiana tabacum* L.

"Promoter" and "promoter region" are terms used interchangeably herein to refer to a DNA sequence that regulates expression of a selected DNA sequence operably linked to the promoter, and which effects expression of the selected DNA sequence in cells. The term also encompasses the 5'untranslated region that may be transcribed into mRNA but is not translated.

"Protein", "polypeptide", and "peptide" are used interchangeably herein when referring to a gene product.

In one aspect, the invention features isolated nucleic acid molecules encoding for PMT1, PMT2, PMT3, PMT4, ADC1, ADC2, ODC, and SAMS, a fragment of NADH dehydrogenase and a fragment of phosphoribosilanthronilate isomerase. The disclosed molecules can be non-coding (e.g. probe, antisense or ribozyme molecules) or can code for a functional enzyme. In one embodiment, the nucleic acid molecules can hybridize to the nucleic acid sequences encoding for PMT1, PMT2, PMT3, PMT4, ADC1, ADC2, ODC, SAMS, a fragment of NADH dehydrogenase, or a fragment of phosphoribosilanthronilate isomerase or to the complements of these nucleic acid sequences. In a preferred embodiment, the hybridization is conducted under mildly stringent or stringent conditions.

In further embodiments, the nucleic acid molecule is at least 50%, 60%, 70%, 80% and more preferably at least 90% or 95% homologous in sequence to the nucleic acid sequences encoding for PMT1, PMT2, PMT3, PMT4, ADC1, ADC2, ODC, SAMS, a fragment of NADH dehydrogenase, or

a fragment of phosphoribosilanthronilate isomerase or to the complements of these nucleic acid sequences. In another embodiment, the nucleic acid encodes a polypeptide that is at least 50%, 60%, 70%, 80% and more preferably at least 90% or 95% similar in sequence to the amino acid sequence of PMT1, PMT2, PMT3, PMT4, ADC1, ADC2, ODC, SAMS, the fragment disclosed herein of NADH dehydrogenase, or the fragment of phosphoribosilanthronilate isomerase disclosed herein.

In another embodiment, the invention features isolated polypeptides, preferably substantially pure preparations, encoded for by the nucleic acid sequences of the invention. Particularly preferred are those polypeptides encoded for by the nucleic acid sequences identified by (SEQ. ID. NO. 2), (SEQ. ID. NO. 5), (SEQ. ID. NO. 8), (SEQ. ID. NO. 11), (SEQ. ID. NO. 13), (SEQ. ID. NO. 15), (SEQ. ID. NO. 18), (SEQ. ID. NO. 21), (SEQ. ID. NO. 23), (SEQ. ID. NO. 25) or (SEQ. ID. NO. 26) or comprising a nucleotide sequence encoding the amino acid sequence encoded by (SEQ ID NO. 3), (SEQ. ID. NO. 6), (SEQ ID. NO. 9), (SEQ. ID. NO. 12), (SEQ. ID. NO. 14), (SEQ. ID. NO. 16), (SEQ. ID. NO. 19), (SEQ. ID. NO. 22) or (SEQ. ID. NO. 24). In particularly preferred embodiments, the subject polypeptides can aid in regulating the production of alkaloids, particularly nicotine, in plants. In one embodiment, the polypeptide is identical to or similar to the protein represented by the amino acid sequences of (SEQ ID NO. 3), (SEQ. ID. NO. 6), (SEQ ID. NO. 9), (SEQ. ID. NO. 12), (SEQ. ID. NO. 14), (SEQ. ID. NO. 16), (SEQ. ID. NO. 19), (SEQ. ID. NO. 22) or (SEQ. ID. NO. 24). In a preferred embodiment, the polypeptide is encoded by a nucleic acid that hybridizes with a nucleic acid represented in.

The polypeptides of the present invention can comprise full length proteins, such as represented by (SEQ ID NO. 3), (SEQ. ID. NO. 6), (SEQ ID. NO. 9), (SEQ. ID. NO. 12), (SEQ. ID. NO. 14), (SEQ. ID. NO. 16), (SEQ. ID. NO. 19), (SEQ. ID. NO. 22) and (SEQ. ID. NO. 24), or can comprise one or more fragments corresponding to one or more particular motifs/domains, or to arbitrary sizes, e.g., at least 5, 10, 25; 50, 100, 150, or 200 amino acids in length.

Another aspect of the invention features chimeric genes comprised of a promoter for the genes for PMT2, PMT1, PMT3, PMT4, or ADC2. Yet another aspect of the invention features chimeric genes or chimeric molecules comprised respectively of the functional gene encoding for or the protein PMT1, PMT2, PMT3, PMT4, ADC1, ADC2, ODC, SAMS, NADH dehydrogenase and/or phosphoribosilanthronilate isomerase.

The invention also concerns isolated and purified promoter regions for tobacco Beta-glucosidase and their use in chimeric genes or chimeric molecules.

Another aspect of the invention involves vectors capable of transporting another nucleic acid to which a vector has been linked. Preferably, the vectors comprise the nucleic acid sequences of the invention or their complements.

The invention also features transgenic plants and their seeds that include (and preferably express) a heterologous form of PMT1, PMT2, PMT3, PMT4, ADC1, ADC2, ODC, SAMS, NADH dehydrogenase and/or phosphoribosilanthronilate isomerase. The present invention also encompasses transgenic plants that contain in their genome a chimeric gene construction 5 incorporating the nucleic acid encoding PMT1, PMT2, PMT3, PMT4, ADC1, ADC2, ODC, SAMS, NADH dehydrogenase and/or phosphoribosilanthronilate isomerase. Such transgenic plants and their seeds may be useful to natively produce enhanced quantities of desirable exogenous proteins, such as compounds useful for pharmaceutical purposes, or proteins that provide herbicide resistance.

Another feature of the invention is the use as probes of the DNA sequences disclosed herein or 10 oligonucleotide fragments thereof. Probes may be useful to obtain additional gene family members or locate homologous genes in tobacco or other plant species. Copies of related genes can be obtained from existing genomic libraries or the genomic libraries can be constructed. In one embodiment, an isolated DNA sequence comprising about a fifteen to about a twenty-five base pair oligonucleotide sequence identical to any consecutive about fifteen to about twenty-five base pair 15 sequence found in the sequences of the invention is used as a probe.

Another feature is use of the polypeptides of the invention in an assay, such as an assay to identify modulators of enzyme activity in plants.

Other features and advantages of the invention will be apparent to those of skill in the art.

The nucleotide and amino acid sequences of the invention are disclosed herein in the Sequence 20 Listing, text, and the figures. Specific sequences of the invention are provided in the attached Sequence Listing and can be understood to represent promoters, nucleic acids, and proteins respectively relating to the following proteins: PMT2 (SEQ. ID. NOS. 1, 2, and 3); PMT1 (SEQ. ID. NOS. 4, 5, and 6); PMT3 (SEQ. ID. NOS. 7, 8, and 9); PMT4 (SEQ. ID. NOS. 10, 11, and 12); SAMS (SEQ. ID. NOS. 13 and 14); ODC (SEQ. ID. NOS. 15 and 16); ADC1 (SEQ. ID. NOS. 17, 25 18, and 19); ADC2 (SEQ. ID. NOS. 20, 21, and 22); ADC1 mRNA (SEQ. ID. NOS. 23 and 24); NADH dehydrogenase (SEQ. ID. NO. 25); and PAI (SEQ. ID. NO. 26). If only two sequence identifiers are provided for a protein, those sequences represent the nucleic acid and the protein respectively. If three identifiers are provided, the identifiers represent promoter, genomic or cDNA nucleic acid, and peptide sequences, respectively. If only one identifier is provided, it represents a 30 DNA fragment coding for the protein or portions of it.

For other reference, the sequences may be found at the following records in GenBank at the following Accession Numbers, which records are hereby incorporated in their entirety herein: AF126810 (NtPMT1); AF126809 (NtPMT2); AF126811 (NtPMT3); AF126812 (NtPMT4), AF176908 (NtomPMT)(*Nicotiana tomentosiformis*); AF76909 (NotoPMT)(*Nicotiana otophora*);

AF127239 (ADC); AF127240 (ADC1); AF127241 (ADC2); AF127242 (ODC); AF233849 (ODC2); AF233850 (ODC1); and AF127243 (SAMS).

The following experimental discussion is presented to better illustrate the invention.

I. PMT

The present invention features the characterization of four members of the nuclear gene family encoding PMT in tobacco *N. tabacum*. The nucleic acid sequences encoding PMT and the amino acid sequences for the proteins are reported herein and can also be found in the DDBJ, EMBL, and GenBank Nucleotide Sequence Databases under the accession numbers for *NtPMT1a*, *NtPMT2*, *NtPMT3*, and *NtPMT4* at AF126810, AF126809, AF126811, and AF126812, respectively. The complete coding region and immediate 5'- and 3'- flanking regions are characterized.

As the discussion below shows, all four PMT genes present in the *N. tabacum* genome are expressed in the roots of wild-type plants and differentially regulated in tobacco lines expressing either high or low total alkaloid contents.

15 Materials and Methods

Plant materials

Seeds of *N. sylvestris*, *N. otophora*, and *N. tomentosiformis* were obtained from the USDA-ARS national tobacco germplasm collection (Oxford, NC). *N. tabacum* cv. Burley 21 and *N. tabacum* cv. Xanthi seeds were kindly provided by Glenn Collins, University of Kentucky. Tobacco plants used for DNA isolation were grown in a soil:vermiculite mixture in the greenhouse under natural lighting conditions. Plants used for RNA extraction were grown in Moltan Plus (Moltan Co., Middleton, TN).

25

Gel blot analysis of genomic DNA

Young leaves were collected from greenhouse grown tobacco (*N. tabacum* cv. Xanthi) plants and total genomic DNA was isolated from freshly-harvested tissues using a modification of the CTAB extraction method (DellaPorta *et al.*, 1983). Approximately 30 µg of total DNA was digested with *Eco*RI, *Kpn*I, or *Eco*RI and *Kpn*I and the digestion products separated by electrophoresis through a 0.75% agarose gel. Restricted and size-fractionated DNA was denatured and transferred to Nytran+ nylon membranes (Schleicher and Schuell, Keene, NH) by capillary blotting in 0.4N NaOH overnight. Membranes were prehybridized in 0.25M Na₂HPO₄ (pH 7.4), 7% SDS, 1 mM Na₂EDTA

for at least 2 hr, then hybridized overnight at 65°C in the same buffer with 2-3 x 10⁶ cpm/mL of a ³²P-labeled single-stranded probe (antisense DNA strand). The probe was prepared by the method of Bednarczuk *et al.* (1991) using a primer (Table 1, primer 4) designed from the 3' end of the *NtPMT1a* coding region (Exon 8) and the full-length coding region of the *NtPMT1a* cDNA as template. The *NtPMT1a* cDNA was generated by RT-PCR using synthetic oligonucleotide primers based on the N- and C-terminal sequences of the A411 cDNA reported by Hibi *et al.* (1994) and RNA template isolated from *N. tabacum* cv. Burley 21 roots. Membranes were washed at a final stringency of 0.1 x SSC, 0.1% SDS at 65°C. Hybridizing bands were visualized by autoradiography and/or imaged using a Molecular Dynamics PhosphorImager (Model 445 SI, Sunnyvale, CA).

10

Genomic library construction and phage isolation

A library of *N. tabacum* cv. Xanthi genomic DNA fragments constructed in EMBL3 was purchased from Clontech (Palo Alto, CA) and a total of 1.1 x 10⁶ recombinant phage were screened by plaque hybridization using random-primed ³²P-labeled *NtPMT1a* cDNA as probe (Sambrook *et al.*, 1989). Prehybridization, hybridization, and washing conditions were as described above. Positive hybridizing phage were plaque purified by subsequent rounds of rescreening and DNA was prepared from 18 independently isolated phage. The phage DNA was characterized by restriction analysis and DNA gel blot analysis and fragments containing the sequences encoding PMT were subcloned into pBluescript KS vectors for further analysis.

Comparison of the hybridizing fragments present in the 18 recombinant phage to the hybridization pattern obtained by genomic DNA blot analysis indicated that only three of the PMT genes suspected to be present in the *N. tabacum* genome were recovered from the library screen. To obtain sequences encoding *NtPMT1a*, a subgenomic library was constructed from *N. tabacum* cv. Xanthi DNA. The library consisted of gel-purified 2.5-3.5 kb *Eco*RI fragments ligated into λ-ZAP II vector arms and packaged using Gigapack III Gold packaging extracts according to the manufacturer's instructions (Stratagene, La Jolla, CA). The primary library was amplified once in *E. coli* XL1-Blue MRF' strain (Stratagene) and screened as described above, except that a random-primed ³²P-labeled *NtPMT1a* cDNA Exon 1-specific probe was used (Table 1). Exon 1 had previously been amplified by PCR using primers 1 and 2 (Table 1) and the *NtPMT1a* cDNA as template. The recombinant phage that hybridized with the probe was isolated from the sublibrary by two more rounds of plaque purification, and the pBluescript phagemid containing the approximate 3.1 kb *Eco*RI genomic fragment with the *NtPMT1a* gene was excised from the λ-ZAP II phage vector using the *in vivo* excision protocol described by Stratagene.

DNA sequence analysis

Unless otherwise noted, DNA sequencing was performed with double-stranded plasmid DNA
5 templates using fluorescent dye terminator technology (dRhodamine Terminator Cycle Sequencing Ready Reaction kit) on an ABI 310 DNA sequencer (Perkin-Elmer Applied Biosystems). For analysis of PCR products, following electrophoretic separation of amplification reaction products, the bands of interest were excised from the polyacrylamide gels, the DNA extracted using the Quiagen Gel Extraction Kit, and the recovered DNA used as sequencing template. Sequencing was
10 performed using AmpliTaq DNA polymerase and fluorescent dye terminator technology (as described above) and primers 1 and 2 (Table 1) specific for Exon 1. Nucleotide and amino acid sequences were analyzed and aligned using either the Clustal method and Lasergene software (DNAStar Inc., Madison, WI) or the PILEUP and ALSCRIPT (Genetics Computer Group, Madison, WI) sequence analysis package (Version 9.0). Transcription factor binding site homologies were
15 identified in promoter DNA sequences by searching the transcription factor database using the GCG program.

RNA gel blot analysis

20 For RNA analysis, roots and other tissues were harvested from mature wild-type (HP; *Nic1*/*Nic2*) and low alkaloid mutant (LP; *nic1*/*nic2*) Burley 21 tobacco plants. For topping experiments, the stem was cut and the top one-third of the plant was removed just prior to flower opening. Roots were harvested just prior to topping (0 hr control) and at various times after decapitation. The tissue was immediately frozen in liquid nitrogen and stored at -80°C until RNA extraction and isolation.
25 Total RNA was isolated from vegetative organs and floral structures of HP and LP Burley 21 tobacco using the TRI-reagent (Molecular Research Center Inc., Cincinnati, OH) and quantified spectrophotometrically by measuring *A*₂₆₀. Total RNA (5 µg) was electrophoresed through 1.2% agarose gels (containing 0.4 M formaldehyde) and transferred to Nytran⁺ nylon membranes. Following prehybridization the membranes were hybridized with a single-stranded *NtPMT1a* cDNA
30 antisense probe (corresponding to the antisense strand of Exons 2 to 8 of the *NtPMT1a* cDNA coding region) as described above. As a control to quantify and normalize RNA levels in each lane, the blot was hybridized with a 400-bp probe derived from the β-ATPase cDNA using primers 6 and 7 (Table 1) as described below.

Semi-quantitative RT-PCR analysis of individual PMT transcript levels

Total RNA (1 µg) extracted from the roots of HP and LP Burley 21 tobacco plants was reverse-transcribed into first-strand cDNA at 42°C using Superscript II reverse transcriptase (Gibco BRL) according to the manufacturer's protocol. Two gene-specific primers were employed in the reactions: primer 5 capable of recognizing Exon 3 of the *PMT* genes and primer 8 specific for Exon 8 of the nuclear gene encoding the β-subunit of mitochondrial ATPase from *N. plumbaginifolia* (*NpATP2.1*) and *N. sylvestris* (*NsATP2.1*) (Boutry and Chua, 1985; Lalanne *et al.*, 1998). The β-ATPase transcript served as an internal reference (constitutively-expressed control) to determine loading accuracy and to normalize expression levels (Kinoshita *et al.*, 1992). Following first strand cDNA synthesis, two sets of nested primers (0.4 µM each primer) were used to amplify the *PMT* and β-ATPase transcripts: primers 1 and 2 (Table 1) recognized Exon 1 in all five *PMT* transcripts and gave products ranging in size from 220 bp to 420 bp and primers 6 and 7 amplified an approximately 400-bp region encompassing a portion of Exons 6 to 8 of the β-ATPase coding region.

Amplification was carried out for 25 cycles using the following reaction conditions: denaturation at 95°C for 1 min, primer annealing at 60°C for 35 sec, and extension at 72°C for 1.5 min; a final extension was conducted at 72°C for 6 min. Amplification products were radioactively labeled by spiking the PCR reaction with 10 µCi 32P-dCTP. Aliquots of the PCR reaction were analyzed on a 6.5% non-denaturing polyacrylamide/1X TBE gel and electrophoresed at 600 volts. The reaction conditions were optimized to provide amplification of both PMT and β-ATPase transcripts in the linear range of the reaction by varying the levels of first strand cDNA template, annealing temperature, and number of cycles of amplification as described in Kinoshita *et al.* (1992). Molecular weight standards were prepared by PCR amplification using the same primers and protocol described above and plasmid DNA templates containing the PMT encoding genomic fragments, as well as genomic DNA from the various *Nicotiana* species indicated in the text.

Following electrophoresis, the polyacrylamide gels were fixed in 5% MeOH, 7.5% acetic acid for 30 min, dried overnight, and used to expose X-ray film. PMT band intensities were quantified using phosphorimager analysis (Molecular Dynamics) and normalized relative to the intensities of the β-ATPase control band in each lane. The experiment was conducted twice with different total RNA samples, and representative results are presented from one of the two experiments.

Results*PMT gene structure and organization in N. tabacum*

Gel blot analysis of total genomic DNA isolated from *N. tabacum* cv. Xanthi, hybridized with a radioactively-labeled cDNA (*NtPMT1a*) encoding the complete coding region of putrescine N-methyltransferase (PMT) showed the presence of five major hybridizing bands in *Kpn*I or *Eco*RI digested DNA, consistent with the presence of a small multigene family in the *N. tabacum* genome (Figure 1).

As part of our initial characterization of the gene family encoding PMT in *N. tabacum*, an EMBL3 genomic library, prepared from *N. tabacum* cv. Xanthi DNA, was screened using the *NtPMT1a* (A411 homologous) cDNA as probe. From a total of 18 recombinant phage isolated, three phage were recovered that contained genomic fragments encoding the *NtPMT2*, *NtPMT3* and *NtPMT4* genes. The three PMT genes were completely encoded within a unique sized *Eco*RI fragment within the phage DNA insert which allowed for the correlation of each with a hybridizing restriction fragment on the gel blot of *N. tabacum* genomic DNA (Figure 1). The complete coding region and immediate 5' and 3' non-coding sequences of the three genes were determined and found to encode full-length PMT proteins (Figure 2). Each PMT gene consisted of 8 exons and 7 introns, consistent with the gene structure reported previously for the PMT genes from *N. sylvestris* (Hashimoto *et al.*, 1998a). Comparison of the deduced amino acid sequences (Figure 2) revealed that the encoded PMT proteins were extremely similar over their entire length, with the only significant variability in primary sequence localized to the extreme N-terminal region of the protein. This region, completely encoded within Exon 1, contains a variable number of an 11 amino acid repeat with a consensus sequence of NGHQNGTSEHQ. The function of the repeated sequence is unknown, but is apparently inconsequential to enzyme function, since the number of repeats does not influence activity and PMTs characterized from other species do not contain the repeated element (Hashimoto *et al.*, 1998a; Suzuki *et al.*, 1999a).

Multiple rounds of screening of the EMBL3 genomic library failed to yield additional hybridizing phage containing sequences encoding the other two PMT genes thought to be present in the *N. tabacum* genome and, therefore, a directed cloning approach was pursued using a subgenomic library constructed from *Eco*RI fragments isolated from *N. tabacum* cv. Xanthi. From this hybridization screening, a phage containing the approximately 3.1 kb *Eco*RI fragment encoding *NtPMT1a* was recovered. The coding region of the *NtPMT1a* gene was found to be identical to the A411 cDNA (Hibi *et al.*, 1994), with the exception of a single base change in Exon 6 that results in a conservative amino acid substitution. This difference could be the result of minor differences among cultivars used in the two studies (i.e., Xanthi vs. Burley 21). Translation of the open reading frame contained in *NtPMT1a* showed that it encoded a protein containing four N-terminal 11 amino acid repeats, similar to Exon 1 of the PMT gene present in *N. tomentosiformis* (Hashimoto *et al.*, 1998a).

Given the observation that *NtPMT1a* encoded a homolog of the *PMT* gene present in *N. tomentosiformis*, the nature and possible evolutionary origin of the remaining *PMT* gene present in the *N. tabacum* genome was brought into question. From our expression studies (described in detail below), we had determined that five distinct *PMT* encoding transcripts were present in the roots of *N. tabacum*, four of which could be accounted for based upon the length of the Exon I coding region in the four *PMT* genes isolated and characterized in our studies described above. The fifth transcript was similar in size to that encoded by *NtPMT1a* and, therefore, was designated *NtPMT1b*. Since the variability in *PMT* gene structure is primarily localized within Exon 1, we used a PCR-based strategy to analyze the *PMT* gene structure and family size in *N. otophora*, the other possible progenitor of *N. tabacum*. As shown in Figure 3, five distinct PCR products were detected in the electrophoretic pattern of amplification products generated from *N. tabacum* genomic DNA using Exon 1 specific primers (Table 1). Consistent with our studies described above and the previous work of Hashimoto *et al.* (1998a), three PCR products were detected in the electrophoretic pattern of amplification products generated from *N. sylvestris* genomic DNA, and a single band was recovered from *N. tomentosiformis* genomic DNA. Amplification of genomic DNA from *N. otophora* using Exon 1 specific primers also yielded only a single band, whose electrophoretic mobility was most similar to that of the *NtPMT1b* derived product.

Analysis of PMT gene intron and flanking sequences

The location of the seven introns within the protein coding region of the five *PMT* genes in *N. tabacum* is identical and appears to be conserved among *PMT* genes from different *Nicotiana* species. There is also little variation in the nucleotide sequences that comprise the Exon-Intron splice junctions in the various *PMT* genes in *N. tabacum* (Table 2). The high degree of nucleotide sequence similarity recognized among *PMT* genes within their coding regions is also present within their introns and immediate 5' and 3' flanking sequences (Table 2 and Figure 4). In general, a greater level of sequence identity is found in the introns of the *NtPMT2*, *NtPMT3*, and *NtPMT4* genes, than in pair-wise comparisons among the introns of the other members of the *N. tabacum* *PMT* gene family. The observed conservation in the intron sequences of the *NtPMT2*, *NtPMT3*, and *NtPMT4* genes is consistent with their origin from the same progenitor species (*N. sylvestris*). One interesting exception occurs within Intron 6, where the length of the intron and the sequence similarity is more conserved between *NtPMT1a* and *NtPMT4*, than between *NtPMT4* and *NtPMT2* or *NtPMT3*.

Approximately 1 kb of nucleotide sequence was determined 5' to the coding regions of the *NtPMT1a*, *NtPMT2*, *NtPMT3*, and *NtPMT4* genes (Figure 4). By comparison to the 5'-untranslated

region (UTR) present in the A411 cDNA, we set the start site for transcription initiation at approximately 57 nucleotides upstream of the MET start codon in *NtPMT1a* and *NtPMT3*, and either 69 or 60 nucleotides upstream in *NtPMT2* and *NtPMT4*. The major distinguishing feature between the 5'-UTRs in the various genes is the presence or absence of a 17 bp sequence in the gene. An appropriately placed TATA-box can be easily recognized 45 bp 5' to the initiation site in all four genes. Within the first 200-250 bp upstream of the TATA box, a high level of sequence conservation is found to exist among the promoter regions in the four genes. After this point, a clear difference can be observed between the *NtPMT1a* promoter and the remaining three genes, and by 400 bp upstream, little similarity can be found among any of the gene family members.

Analyzing the proximal regions of the various *PMT* promoters with various motif scanning software identified several G-box-like sequences (Foster *et al.*, 1994; Kim *et al.*, 1992; Menkens *et al.*, 1995; Staiger *et al.*, 1989; Williams *et al.*, 1992) at various positions among the *PMT* promoters, and a potential metal response element (MRE) (positions -75 to -66; numbering relative to the *NtPMT1a* promoter sequence) in three of the four *PMTs* (Cizewski-Culotta and Hamer, 1989; Thiele, 1992). An unusual 17 nucleotide stretch of guanine occurs at positions -259 to -243 in the *NtPMT1a* gene promoter followed upstream by a purine-rich region (positions -332 to -263). In the *NtPMT3* promoter a 14 bp palindromic sequence (positions -497 to -484) was detected. *PMT* gene expression has been reported to increase in root tissues following treatment with methyl jasmonate (Imanishi *et al.*, 1998). However, none of the sequence motifs reported to confer methyl jasmonate-responsiveness in other plant genes (Mason *et al.*, 1993; Rouster *et al.*, 1997) were detected in the *PMT* promoters.

Comparison of the available nucleotide sequence information from the 3'-flanking regions of the various *PMT* genes in *N. tabacum* revealed that the 3'-UTRs in the *NtPMT2*, *NtPMT3*, and *NtPMT4* genes of *N. tabacum* share approximately 81-94% identity with each other and are essentially identical to those reported for *N. sylvestris* PMTs by Hashimoto *et al.* (1998a). The major distinguishing feature among the various genes is the presence of two short (20 bp and 4 bp) deletions in the *NtPMT2* gene, which lowers the percent identity. The 3'-UTR of *NtPMT1a* is identical to that reported for the A411 cDNA (Hibi *et al.*, 1994) and 81-94% identical to the other *PMT* genes in the *N. tabacum* genome. Unfortunately, no sequence information is currently available for the 3'-UTR of the *N. otophora* or *N. tomentosiformis* *PMT* genes.

Regulation of PMT gene expression

To determine whether the members of the *PMT* gene family in *N. tabacum* were differentially

expressed, a series of experiments were carried out to define the temporal and spatial distribution of transcripts arising from the five genes. Shown in Figure 5A are the results of gel blot analysis of total RNA extracted from various tissues of mature Burley 21 tobacco plants hybridized with radioactively-labeled probe capable of detecting all five *PMT* transcripts. Consistent with previous studies (Hashimoto *et al.*, 1998b; Hibi *et al.*, 1994), *PMT* expression is localized exclusively to roots. When maturing wild-type (HP) Burley 21 plants are topped (i.e., the floral meristem and upper 1/3 of the stem are removed), a dramatic increase in *PMT* transcript abundance is observed within 2 hr, reaching a maximal level of accumulation by 12-24 hr. Two size transcripts are detected on the gel blots, reflecting the small difference in message size that occurs as a result of the difference in size of Exon 1 among the genes.

In addition to examining *PMT* gene expression in wild-type plants, we also examined expression in a low nicotine-producing (LP) mutant of Burley 21 (Legg and Collins, 1971). The low nicotine Burley 21 line harbors mutations at two independent loci (*nic1* and *nic2*) thought to be global regulators of gene expression involved in alkaloid formation. As shown in Figure 6B, topping of the low nicotine mutant (*nic1nic2*) Burley 21 did not cause an increase in *PMT* transcript abundance as observed in wild type plants. Thus, it appears that *Nic1* and *Nic2* are likely involved in regulation of *PMT* expression in the very least, and may also be involved in the regulation of other genes in the alkaloid biosynthetic pathway. Whether this is a direct effect (e.g., transcriptional activation) or indirect remains to be determined.

In order to determine the extent to which the individual members of the gene family contributed to the general pattern of expression described above, a semi-quantitative RT-PCR strategy (Kinoshita *et al.*, 1992) was used to detect and quantify the levels of the individual *PMT* transcripts in the roots of both wild-type (HP) and low alkaloid (LP) Burley 21 tobacco. This approach takes advantage of the fact that Exon 1 is variable in length among the various *PMT* genes (Figure 2), allowing for their individual detection and quantitation following polyacrylamide gel electrophoresis and autoradiography.

Five RT-PCR products (representing Exon 1 from each of the five genes present in *N. tabacum*) were detected in the electrophoretic profiles of amplification products derived from reactions using either HP or LP Burley 21 root RNA (Figure 6A). All five *PMT* genes present in the *N. tabacum* genome were expressed in the roots of wild-type plants, and topping resulted in a differential accumulation of transcripts derived from each gene. Among the five genes, transcripts derived from the *NtPMT2* and *NtPMT1b* showed the greatest increase in abundance rising approximately 3-fold during the first 24 hr post-topping, whereas levels of the *NtPMT1a* and *NtPMT4* transcripts changed little in response to topping (Figure 6B). In the LP mutant, little or no effect was observed on the

levels of the various *PMT* transcripts following topping, although in some cases (e.g., *NtPMT1a*) a small but likely insignificant depression in transcript abundance was detected. Thus, it appears that all five genes contribute to PMT activity levels within the root.

5 II. ADC

The present invention features the characterization of two members of the nuclear gene family encoding ADC in tobacco *N. tabacum* L. As the following discussion shows, *ADC2* is preferentially expressed in roots and accounts for the major portion of *ADC* transcripts present. Furthermore, analysis of *ADC* transcript levels in roots of low and high nicotine producing lines showed that *ADC* expression is under the control of the *Nic1 Nic2* regulatory loci.

Materials and methods

Plant growth and tissue preparation

15 Seeds of *N. tabacum* cv. Xanthi, wild-type and low alkaloid *nic1 nic2* mutant *N. tabacum* cv. Burley 21 were obtained from Dr. G. Collins (University of Kentucky, Lexington). Tobacco plants used for DNA isolation were grown in soil:vermiculite mixture in the greenhouse under natural lighting conditions. Plants used for RNA extraction were grown either in Moltan Plus (Moltan Co.,
20 Middleton, TN) or hydroponically in a dilute (half-strength) Peters nutrient solution with continuous aeration of the roots under natural lighting conditions in the greenhouse. Topping experiments were conducted by removing the floral meristem, leaves and stem (approximately the upper 1/3 of the plant) from tobacco plants just prior to blooming. Plant tissues were collected from fully matured individuals, frozen in liquid nitrogen, and stored at -80°C until used for RNA preparation (see
25 below).

Screening of genomic libraries and phage characterization

A genomic library constructed in λ EMBL3 from *N. tabacum* cv. Xanthi leaf DNA (Clonetech, Inc.,
30 Palo Alto, CA) was screened by plaque hybridization (Sambrook *et al.*, 1989) using an [α -³²P]-dCTP-labeled, 2.7 kb *EcoRI-XbaI* fragment from plasmid PR24 as probe. PR24 encodes a full length ADC cDNA isolated from the roots of wild-type *N. tabacum* cv. Burley 21 (Wang, 1999). Hybridization was performed at 65 °C for 16 h in a solution containing 0.25 M Na₂HPO₄ (pH 7.2) and 7% (w/v) SDS. Following hybridization, the membranes were washed twice in 2 x SSC, 0.1%

SDS for 15 min at room temperature, once in 0.2 x SSC, 0.1% SDS for 30 min at 65°C. Hybridizing phage were picked and plaque purified through three subsequent rounds of hybridization screening. Phage DNA was isolated from plaque purified phage using a Qiagen Phage Midi Preparation Kit (Qiagen, Germany) and insert DNA characterized by restriction mapping and DNA gel blot analysis.

5 The relevant hybridizing bands in each phage were cloned into pBluescript SK+ vectors for further analysis.

Nucleic acid sequencing and analysis

10 Nucleotide sequencing was carried out manually using the Sequenase Version 2.0 protocols according to the manufacturer's protocol (United States Biochemical, Cleveland, OH) or with an ABI 310 Genetic Analyzer (PE Applied Biosystems, Foster City, CA) using double-stranded plasmid DNA templates prepared utilizing the Qiaprep Spin Plasmid Kit (Qiagen USA, Valencia, CA). The nucleotide and predicted amino acid sequences of the various cDNAs were analyzed using BLAST sequence analysis programs (Altschul *et al.*, 1990; Gish and States, 1993) and protein sequence alignments were carried out using the PILEUP program (Genetics Computer Group Sequence Analysis package, Version 9.0 (GCG, University of Wisconsin, Madison, WI) and the various gene sequences available in the NCBI (National Center for Biotechnology Information, Bethesda, MD) nucleotide and protein sequence database. Manual adjustment of the sequence alignments were carried out as necessary.

15

20

RNA isolation and gel blot analysis

Total RNA was extracted from tobacco roots, leaves, and floral parts using Tri-Reagent (Molecular Research Center, USA, Cincinnati, OH) according to the manufacturer's protocol. For RNA gel blot analysis, aliquots (10 µg) of total RNA extracted from the various tissues were fractionated by electrophoresis through a 1.2% agarose-formaldehyde gel and blotted onto Nytran nylon membranes (Schleicher & Schuell, Keene, NH) using 10 X SSC. The transferred RNA was UV cross-linked to the membrane using a UV Stratalinker (Stratagene, La Jolla, CA) and the membranes were prehybridized in 7% SDS, 0.25 M Na₂HPO₄, pH 7.2 for 2-4 hours at 65°C.

25

30

Hybridization was carried out in the same buffer in the presence of ³²P-labeled probes for 16 hr at 65°C. The membranes were washed under high stringency conditions and subject to autoradiography at -80°C for approximately 48 h.

For gel blot analysis, [α - ³²P]-dCTP -labeled probes were prepared by random primed labeling

(Random Primed Labeling Kit, Boehringer Mannheim, Indianapolis, IN) using 25-50 ng of a 2.7 kb *Eco*RI-*Xba*I fragment derived from PR24 and a 460 bp fragment amplified from the β - subunit of the tobacco mitochondrial ATP synthase gene (*atp2*) (Boutry and Chua, 1985).

5 *Semi-quantitative RT-PCR analysis of NtADC1 and NtADC2 transcript levels.*

Total RNA (2 μ g) from roots, leaves, or floral parts was reverse transcribe at 40°C for 1 h in a reaction cocktail containing 200 units of SuperscriptII reverse transcriptase (RNase H-, Gibco BRL, USA), 10 units RNase inhibitor (Perkin Elmer), 200 μ m dNTPs and 40 pmol of primer, in total 10 volume of 20 μ l. For first strand cDNA synthesis, a single primer [5'-AGAAAAACATCACCAACT-3'] capable of hybridizing to both the *ADC1* and *ADC2* transcripts was used in the reaction. As a control, a primer (5'-GCAACTGTCATCTTATCATCTTC-3') specific for the β -subunit of the tobacco mitochondrial ATP synthase gene *atp2* (Boutry and Chua, 1985) was used in the reverse transcriptase reaction.

15 Following reverse transcription, the single stranded cDNA products were serially diluted over a concentration range between 1 to 50 ng RNA, and PCR amplification was carried out for 25 cycles of 45 s at 94°C, 1 min at 64°C and 1 min at 72°C in a Genemate thermocycler (ISC Bioexpress, UT). The reaction mixture contained cDNA template, 1 x PCR buffer (Boehringer Mannheim), 100 μ M dNTPs, 25 pmol of each forward and reverse primer and 1 unit Taq DNA polymerase. The PCR reactions specific for *ADC1* transcripts contained the following primers: ADC1-forward, 5'-CGTAGACGCTACTGTTTC-3' and ADC1-reverse, 5'-TGGACAAC TGTGGAGGCG-3'. Reactions specific for *ADC2* transcripts contained primers ADC2-forward, 5'-TGTAGATGCTGCTGTTGTT-3', and ADC2-reverse, 5'-TGAACAAAC TGCGGAGGCA-3'. Control reactions for normalization of amplification products contained 25 pmol of primers specific 20 for the tobacco *atp2* transcripts: *atp2* forward, 5'-GTATATGGTCAAATGAATGAGCC-3', and *atp2* reverse.int, 5'-GCAGTATTGTAGTGATCCTCTCC-3'. For quantitation purposes, amplification reactions were supplemented with 1 μ Ci 32 P-dCTP. PCR products were separated by electrophoresis through 1.2% agarose gels, the fractionated reaction products transferred onto a Hybond N+ membranes, dried and subject to autoradiography at -70°C. Quantitation was carried out by 25 phosphorimaging using a Molecular Dynamics PhosphorImager. Values were normalized relative to the intensities of the *atp2* control band in each lane. The experiment was conducted twice with different total RNA samples, and representative results are presented from one of the two experiments.

Results

These studies show the structure and expression of individual members of the *ADC* gene family in tobacco. An α -³²P-dCTP-labeled 2.7 kb EcoRI-XbaI fragment from PR24 encoding the 5 ADC coding region was used to screen an λ EMBL3 phage genomic library. From a screen of approximately 3×10^5 phage, seventeen hybridizing phage were recovered, of which five were fully characterized by restriction mapping and DNA gel blot analysis. These phage fell into two groups based on their restriction profile. The relevant hybridizing fragments from the various phage were cloned into pBluescript and their nucleotide sequence determined.

Presented in Figure 7 are the nucleotide and predicted amino acid sequences of NtADC-1 and NtADC-2 genes. Both genes contain a single open reading frame, uninterrupted by introns. The nucleotide and amino acid sequence encoded in NtADC-1 is identical to that of PR24, the full length cDNA isolated from *N. tabacum* cv Burley 21. There are 84 nucleotide differences within the NtADC-1 and NtADC-2 coding regions, resulting in 23 amino acid differences between the ADC1 and ADC2 proteins, respectively. The ADC1 protein is one amino acid shorter in length, missing Val-13.

By comparison to the full-length cDNA, the 5'-untranslated region (UTR) present in NtADC-1 and NtADC-2 are 431 bp and 432 bp long, respectively. The size of the 5'-UTR in the ADC transcripts is considerably larger than the average size of the plant leader sequence (Joshi, 1987). In contrast, the 3'-UTRs present in NtADC-1 and NtADC-2 are relatively short, approximately 84 nucleotides in length. In both gene sequences, a conserved polyadenylation signal (AATAATA) can be recognized 23 nucleotides from the site of polyadenylation site found in the PR24 cDNA.

Pairwise comparison of the *N. tabacum* ADC1 and ADC2 proteins with the ADCs of other plant species showed that the *N. tabacum* proteins are approximately 82% identical to the ADC of its evolutionary progenitor species *N. sylvestris* [Genbank Accession No. AB012873] and 86% identical to the ADC from tomato (*Lycopersicon esculentum*) [31], another member of the Solanaceae family (Figure 2). As might be expected, the *N. tabacum* ADC shares considerably less similarity to ADCs isolated from species more distantly related evolutionarily, such as *Arabidopsis* - 67% identical [32, 33], soybean- 67% identical [34], and oat - 42% identical [35] and is only 29% identical to the enzyme from *Escherichia coli* - [36].

The predicted protein coding regions for the *N. tabacum* ADCs are substantially longer than those reported for the ADC proteins of *N. sylvestris* and *L. esculentum* [31], but are similar in length to those reported in *Arabidopsis*, oat, soybean [32-35] and for the *E. coli* enzyme [36]. The

difference in overall length appears to arise from an apparent nucleotide deletion in the *N. sylvestris* and tomato cDNA sequences relative to the ADC1 and ADC2 predicted sequence and those in other plants. In the nucleotide sequences reported for both the *N. sylvestris* and tomato cDNAs, a guanine residue (position 2295 in the *N. sylvestris* sequence and 1531 in the tomato sequence) is missing [Genbank Accession No. AB012873]. This deletion changes the reading frame and introduces a premature termination to the predicted coding region. Using the sequence information available in the NCBI database, correcting for this error allowed us to extend the predicted C-terminus of the both ADC proteins, yielding the alignment to the *N. tabacum* ADCs and those of other plant ADCs as indicated in Figure 8. We have also included in the alignment shown in Figure 8, the correction at the N-terminus of the predicted tomato ADC protein sequence noted by Pérez-Amado et al. [37], allowing better alignment of all of the higher plant sequences.

Developmental regulation of arginine decarboxylase expression

It has been shown that nicotine formation can be activated in the roots of maturing tobacco plants by topping, that is, removal of the flower head and several young leaves (Akehurst, 1981; Hibi, et al., 1994). Coincident with the activation of nicotine formation, there is an increase in the levels of transcripts encoding ODC, PMT and spermidine synthase (SPS) over the subsequent 24 hr period in wild-type plants (Hibi et al., 1994; Riechers and Timko, 1999). To determine the effects of topping on *ADC* expression in roots, Burley 21 plants were grown in the greenhouse to the bud stage at which point the upper 1/3 of the plant was removed and samples of roots tissues were collected before and at various times post-topping. As shown in Figure 9, *ADC* message abundance increased in the roots of topped Burley 21 plants during the 24 hr period after topping. Low alkaloid (LA) mutants of Burley 21 show a much lower level of *ADC* expression in their roots, and no induction of *ADC* transcript accumulation after topping. The lack of *ADC* induction in the low-alkaloid mutant is consistent with previous studies (Hibi et al., 1994; Riechers and Timko, 1999; Wang, 1999) showing a general inability to activate gene expression leading to increased polyamine formation and alkaloid biosynthesis as a result of the mutation of the *Nic1* and *Nic2* regulatory genes.

NtADC-2 is predominately expressed in roots of wild-type plants.

Due to the high degree of identity between the NtADC-1 and NtADC-2 transcripts (e.g., 95.8% coding regions, 94.4% and 96.4% in 5'- and 3'-UTRs, respectively), it is impossible to distinguish between the two transcripts by RNA gel blot analysis. Therefore, we employed a RT-PCR based

strategy and gene specific oligonucleotide primers. Total RNA was extracted from tobacco roots, leaves and flowers, and single-stranded cDNA synthesized using an oligonucleotide primer capable of hybridizing to both ADC1 and ADC2 transcripts. As an internal control for amplification, a gene specific primer recognizing the *atp2* transcript encoding the β -subunit of the tobacco mitochondrial ATPase was included in the reactions. Under experimental conditions providing amplification in the linear range of the PCR reaction, gene specific forward and reverse primers were used to specifically amplify either ADC1 or ADC2 cDNAs. Test reactions (Figure 10A) using plasmid DNA encoding NtADC1 or NtADC2 as template demonstrated the specificity of the primers. As shown in Figure 10B, the main transcripts detectable in all tissues tested are derived from NtADC-2. Flowers express the highest level of ADC, and leaves lowest. In the flowers, although ADC1 is detectable, far less than ADC2 Roots also express a significant level of ADC.

ADC transcript levels are highest in the roots and floral organs, and low in other plant tissues. The two ADC genes investigated appear to have different modes of regulation, with ADC2 being predominately expressed in the roots and other organs.

At the present time, only limited information is available on the nature of regulatory regions in the promoters of genes encoding enzymes of alkaloid biosynthesis. The availability of cloned genomic fragments encoding ADC allows one to begin mapping regulatory sequences within members of these genes responsible for tissue specific, developmental, and inducible expression.

20

III. ODC

The present invention features the genes of two members of the nuclear gene family encoding ODC in tobacco *N. tabacum*. As the following experimental discussion shows, the ODC-2 gene is preferentially expressed in roots and floral tissues. Furthermore, the abundance of ODC transcripts in root tissues is affected by topping. Furthermore, analysis of ODC transcript levels in roots of low and high nicotine producing lines shows that ODC expression is under the control of the *Nic1 Nic2* regulatory loci.

Materials and methods

30

Plant growth and tissue preparation

Seeds of *N. tabacum* cv. Xanthi, wild-type and low alkaloid *nic1 nic2* mutant *N. tabacum* cv. Burley 21 were obtained from Dr. G. Collins (University of Kentucky, Lexington). Tobacco plants used for DNA isolation were grown in soil:vermiculite mixture in the greenhouse under natural lighting

conditions. Plants used for RNA extraction were grown either in Molton Plus (Molton Co., Middleton, TN) or hydroponically in a dilute (half-strength) Peters nutrient solution with continuous aeration of the roots under natural lighting conditions in the greenhouse. Topping experiments were conducted by removing the floral meristem, leaves and stem (approximately the upper 1/3 of the plant) from tobacco plants just prior to blooming. Floral parts and other plant tissues were collected from fully matured individuals, frozen in liquid nitrogen, and stored at -80°C until used for RNA preparation (see below).

Screening of genomic libraries and phage characterization

A genomic library constructed in EMBL3 from *N. tabacum* cv. Xanthi leaf DNA (Clonetech, Inc., Palo Alto, CA) was screened by plaque hybridization (Sambrook *et al.*, 1989) using a ³²P-radiolabeled, 1.6 kb *EcoRI-XbaI* insert from plasmid PR46 as probe. PR46 encodes a full length ODC cDNA previously isolated by differential screening of plasmid libraries prepared from mRNA isolated from the roots of wild-type Burley 21 plants before and 3-days post-topping (Wang, J., Sheehan, M., Bookman, H. and Timko, M.P., unpublished data). Hybridization was performed at 65°C for 16 h in a solution containing 0.25 M Na₂HPO₄ (pH 7.2) and 7% (w/v) SDS. Following hybridization, the membranes were washed twice in 2 x SSC, 0.1% SDS for 15 min at room temperature, once in 0.2 x SSC, 0.1% SDS for 30 min at 65°C. Hybridizing phage were picked and plaque purified through three subsequent rounds of hybridization screening. Phage DNA was isolated from plaque purified phage using a Qiagen Phage Midi Preparation Kit (Qiagen USA, Valencia, CA) and insert DNA characterized by restriction mapping and DNA gel blot analysis. The relevant hybridizing bands in each phage were cloned into pBluescript SK+ vectors for further analysis.

Nucleic acid sequencing and analysis

Nucleotide sequencing was carried out manually using the Sequenase Version 2.0 protocols according to the manufacturer's protocol (United States Biochemical, Cleveland, OH) or with an ABI 310 Genetic Analyzer (PE Applied Biosystems, Foster City, CA) using double-stranded plasmid DNA templates prepared utilizing the Qiaprep Spin Plasmid Kit (Qiagen USA, Valencia, CA). The nucleotide and predicted amino acid sequences of the various cDNAs were analyzed using BLAST sequence analysis programs (Altschul *et al.*, 1990; Gish and States, 1993) and protein sequence alignments were carried out using the PILEUP program (Genetics Computer Group Sequence Analysis package, Version 9.0 (GCG, University of Wisconsin, Madison, WI) and the various gene sequences available in the NCBI (National Center for Biotechnology Information, Bethesda, MD) nucleotide and protein sequence database. Manual adjustment of the sequence alignments were

carried out as necessary.

RNA isolation and gel blot analysis

Total RNA was extracted from tobacco roots, leaves, and floral parts using Tri-Reagent (Molecular Research Center, USA, Cincinnati, OH) according to the manufacturer's protocol. For RNA gel blot analysis, aliquots (10 µg) of total RNA extracted from the various tissues were fractionated by electrophoresis through a 1.2% agarose-formaldehyde gel and blotted onto Nytran nylon membranes (Schleicher & Schuell, Keene, NH) using 10 X SSC. The transferred RNA was UV cross-linked to the membrane using a UV Stratalinker (Stratagene, La Jolla, CA) and the membranes were prehybridized in 7% SDS, 0.25 M Na₂HPO₄, pH 7.2 for 2-4 hours at 65°C. Hybridization was carried out in the same buffer in the presence of ³²P-labeled probes for 16 hr at 65°C. The membranes were washed under high stringency conditions and subject to autoradiography at -80°C for approximately 48 h.

Restriction fragments derived from cDNA clones of interest were separated by agarose gel electrophoresis, the DNA was purified, and quantified by spectrophotometry. [³²P]-dCTP-labeled probes were prepared from 25-50 ng of insert DNA by random primed labeling (Random Primed Labeling Kit, Boehringer Mannheim, Indianapolis, IN). As a control, the blots were also probed with radioactively labeled probes encoding the alkaloid biosynthesis enzyme putrescine N-methyltransferase (PMT) (Riechers and Timko, 1999), a root specific, topping inducible β-glucosidase encoding cDNA (TBG-1) (Riechers, D.E. and Timko, M.P.; unpublished data), 26S rRNA (PR31) or 28S rRNA fragments.

Genomic DNA isolation and gel blot analysis

Tobacco genomic DNA was prepared from tobacco leaf tissue by the method of Junghans and Metzlaff (1990). Total genomic DNA (15 µg) was digested to completion with *Eco*RI or *Hind*III, the digestion products were fractionated by electrophoresis through a 0.8% (w/v) agarose gel, and transferred onto Nytran nylon membrane (Schleicher & Schuell, Keene, NH) in the presence of 0.4 N NaOH (Sambrook *et al.*, 1989). Following transfer, the membrane was rinsed in 2 X SSC, the DNA was UV cross-linked to the membrane, and the membrane was prehybridized and hybridized as described above. Following hybridization and washing, the membranes were subjected to autoradiography at -80°C.

Results and discussion

Gel blot analysis of tobacco genomic DNA cut with various restriction enzymes and hybridized with an [α -³²P]-dCTP-labeled 1.6 kb *EcoRI-XbaI* cDNA fragment (PR46) encoding the full-length ODC protein from *N. tabacum* cv Burley 21 (Wang, J., Sheehan, M., Bookman, H. and Timko, M.P., unpublished data) indicated ODC is encoded by small gene family in the *N. tabacum* genome (Fig. 11). Four to five major bands and several minor bands of sufficient size to encode full-length genes are detected in either *EcoRI* or *HindIII* digested tobacco DNA.

To further analyze the structure and regulation of members of the *ODC* gene family in tobacco, a λ EMBL3 phage genomic library constructed with DNA from *N. tabacum* cv Xanthi was screened using a [α -³²P]-labeled probes prepared from PR46 (as described above). From a screen of approximately 3×10^5 phage, five hybridizing phage were recovered, of which three were fully characterized by restriction mapping and DNA gel blot analysis. Two phage proved to contain identical insert DNA and the third had a unique restriction digestion profile. Following DNA gel blot analysis, the hybridizing fragments were cloned into pBluescript and their nucleotide sequence determined.

The complete *NtODC-2* gene spans two *SaII* fragments of 2.7 kb and 6.5 kb. The coding region of the gene contains a single 1302 bp open reading frame uninterrupted by introns (Fig. 12). The nucleotide sequences of *NtODC-2* is identical within the coding and 5' and 3'- untranslated regions to the PR46 encoded cDNA, with the exception of four nucleotide changes (residues +2, +4, +6 and +8) in the 5'-untranslated region. These nucleotide differences likely reflect changes introduced during the cDNA synthesis reaction.

The predicted amino acid sequence for the *NtODC-2* encoded protein (designated pODC2) (Fig. 13) is identical to the ODC characterized from Burley 21 tobacco encoded by PR46 (Wang, J., Sheehan, M., Bookman, H. and Timko, M.P., unpublished data) and to the partial *N. tabacum* ODC cDNA sequence (PR17) reported by Malik *et al.*, (1996). Comparison of the predicted amino acid sequence for pODC2 with the ODC proteins characterized from two different tobacco cultivars showed that the pODC2 differs by 7 amino acid (98% identity) from the ODC protein characterized from the high alkaloid cultivar, *N. tabacum* cv. SC58 [Genbank Accession No. Y10472.1] and by 7 amino acid (98% identity) from ODC protein from BY-2 cells. The tobacco pODC2 is 89% and 90% identical to the ODCs from tomato (*Lycopersicon esculentum*) and jimsonweed (*Datura stramonium*), respectively, but substantially less similar to ODCs from yeast (35% identity) and humans (32% identity).

The *NtODC-1* gene, contained on an 4.0 kb *XbaI* fragment, encodes a single open reading frame of 141 amino acids encompassing the amino terminal one-half of ODC (Fig. 12). Six amino acid residue changes distinguish the *NtODC-2* and *NtODC-1* encoded proteins over the homologous

region of the proteins. Beginning at amino acid residue 130, the *NtODC-1* encoded protein (pODC1) diverges from pODC2, with a stop codon present after residue 141. Scanning the available nucleotide sequence (> 1 kb) in the 3'-flanking region of the *NtODC-1* gene failed to reveal any evidence for ODC homologous protein sequences in any of the three translational reading frames.

5 Interestingly, a comparison of the 5'-flanking sequence of the *NtODC-1* and *NtODC-2* genes revealed that while the *NtODC-2* gene has a clearly recognizable TATA-box properly located at approximately -35 bp from the transcriptional start site, no such regulatory motif is found in the *NtODC-1* gene sequence. Consistent with this observation, RNA gel blot analysis performed using a hybridization probe prepared from *NtODC-1* immediately downstream of the frame shift, failed to
10 detect any message in various tissues of mature tobacco plants (data not shown). Thus, it appears that *NtODC-2* represents an unexpressed pseudogene in the *N. tabacum* genome.

To determine the spatial pattern of expression of the *NtODC-2* gene, gel blot analysis was carried out using total RNA prepared from roots, stems, young and mature leaves, and various floral parts of Burley 21 tobacco plants. As shown in Fig 14, transcripts encoding ODC were easily
15 detected in the roots, with little or no expression in other tissues except sepals, carpels, and mature stamens.

The formation of nicotine and total leaf alkaloids in tobacco is known to be under the control of at least two independent genetic loci (Legg *et al.*, 1969; Legg and Collins, 1971), designated *Nic1* and *Nic2* (Hibi *et al.*, 1994). *Nic1* and *Nic2* are semidominant and operate synergistically to control
20 plant alkaloid content, with mutations within these genes resulting in plants with reduced levels of nicotine and total leaf alkaloids (wild-type > *nic1* > *nic2* > *nic1 nic2*) (Legg *et al.*, 1969; Legg and Collins, 1971). Although no information is available on the nature of their encoded products, it has been speculated that *Nic1* and *Nic2* likely encode transcriptional regulators capable of globally interacting with a subset of genes encoding components of polyamine and alkaloid biosynthesis
25 (Hibi *et al.*, 1994). Removal of the flower head and several young leaves (i.e., topping) leads to activation of nicotine formation in the roots of decapitated plants (Akehurst, 1981; Hibi *et al.*, 1994). To determine the effects of topping on *NtODC-1* expression in roots, Burley 21 plants were grown in the greenhouse to the bud stage at which point the upper 1/3 of the plant was removed and samples
30 of roots tissues were collected before and at various times post-topping. As shown in Fig 14B, low levels of the *ODC* transcripts were found in roots prior to topping and message abundance increased approximately 2-fold in the roots of topped Burley 21 plants 4 hr after topping. By 24 hr after topping, *ODC* transcript levels return to their initial levels. Low alkaloid mutants of Burley 21 subjected to the same treatment show a much lower level of stimulation of *ODC* transcript accumulation after topping, and the enhanced transcript abundance does not persist beyond 4 hr. By

comparison, transcripts encoding PMT and and a tobacco root-specific β -glucosidase (TBG-1) show patterns of accumulation similar to that observed for ODC transcripts in wild-type plants, but no induction in the low-alkaloid mutant, consistent with previous studies (Hibi *et al.*, 1994; Riechers and Timko, 1999; Wang, 1999).

5

IV. SAMS

A single recombinant phage is identified as encoding for SAMS. This λ phage contains an approximately 15kB Sall insert. Restriction mapping and PCR analysis indicates that the insert DNA contains primarily the coding and 3'non-coding portions of the SAMS gene. The nucleotide sequences for the gene encoding SAMS can be found at GenBank Accession Nos. AF27243 (full length SAMS cDNA).

V. NADH dehydrogenase

A fragment of the cDNA encoding for NADH dehydrogenase in *N. tabacum* shows high expression in the roots of mature wild-type HP plants compared to low alkaloid mutant LP plants.

VI. Phosphoribosylanthranilate isomerase (PAI)

The gene encoding for a fragment of phosphoribosylanthranilate isomerase in *N. tabacum* is a homolog of the *Arabidopsis thaliana* gene encoding PAI, an enzyme involved in tryptophan biosynthesis. This enzyme is involved in the overall formation of aromatic compounds in plants.

REFERENCES

Akehurst BC. 1981. The growth, plant structure and genetics. In: Rhind D, Wrigley G, eds., Tobacco, London: Longman Press, 45-95.

25

Alabadi D, Carbonell J. 1998. Expression of ornithine decarboxylase is transiently increased by pollination, 2,4-dichlorophenoxyacetic acid, and gibberellic acid in tomato ovaries. *Plant Physiology* 118: 323-328.

30

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic alignment search tool. *Journal of Molecular Biology* 215: 403-410.

Baldwin IT. 1989. Mechanism of damage-induced alkaloid production in wild tobacco. *Journal of Chemical Ecology* 15: 1661-1680.

Baldwin IT, Prestin CA. 1999. The eco-physiological complexity of plant responses to insect herbivores. *Planta* 208:137-145.

5 Baldwin IT, Schmelz EA, Ohnmeiss TE. 1994. Wound-induced changes in root and shoot jasmonic acid pools correlate with induced nicotine synthesis in *Nicotiana sylvestris* Spegazzini and Comes. *Journal of Chemical Ecology* 20: 2139-2157.

Baldwin IT, Schmelz EA, Zhang Z-P. 1996. Effects of octadecanoic metabolites and inhibitors on induced nicotine accumulation in *Nicotiana sylvestris*. *Journal of Chemical Ecology* 22: 61-74.

10 Baldwin IT, Zhang Z-P, Diab N, Ohnmeiss TE, McCloud ES, Lynds GY, Schmelz EA. 1997. Quantification, correlations, and manipulations of wound-induced changes in jasmonic acid and nicotine in *Nicotiana sylvestris*. *Planta* 201: 397-404.

15 Bell E. and R.L. Malmberg, Analysis of a cDNA encoding arginine decarboxylase from oat reveals similarity to the *Escherichia coli* arginine decarboxylase and evidence of protein processing. *Mol. Gen. Genet.*, 224 (1990) 431-436.

20 Boutry M. and N.H. Chua, A nuclear gene encoding the beta subunit of the mitochondrial ATP synthase in *Nicotiana plumbaginifolia*. *EMBO J.*, 4 (1985) 2159-2165.

Bracher D, Kutchan TM. 1992. Strictosidine synthase from *Rauvolfia serpentina*: analysis of a gene involved in indole alkaloid biosynthesis. *Archives of Biochemistry and Biophysics* 294: 717-723.

25 Chattopadhyay MK, Ghosh B. 1998. Molecular analysis of polyamine biosynthesis in higher plants. *Current Science* 74, 517-522.

Chou W-M, Kutchan TM. 1998. Enzymatic oxidations in the biosynthesis of complex alkaloids. *Plant Journal* 15, 289-300.

30 De Luca V. and B. St. Pierre. 2000. The cell and developmental biology of alkaloid biosynthesis. *Trends in Plant Science* 5: 168-173.

Eilbert U. 1998. Induction of alkaloid biosynthesis and accumulation in plants and *in vitro* cultures

in response to elicitation. In: Roberts MF, Wink M, eds. *Alkaloids: Biochemistry, Ecology, and Medicinal Applications*. New York: Plenum Press, 219-262.

5 **Facchini PJ, Penzes-Yost C, Samanani N, Kowalchuk B.** 1998. Expression patterns conferred by tyrosine/dihydroxyphenylalanine decarboxylase promoters from opium poppy are conserved in transgenic tobacco. *Plant Physiology* 118: 69-81.

Galloway G.L., R.L. Malmberg and R.A. Price, Phylogenetic utility of the nuclear gene arginine decarboxylase: an example from Brassicaceae. *Molec. Biol. & Evol.*, 15 (1998) 1312-1320.

10 **Gantet P, Imbault N, Thiersoult M, Doireau P.** 1998. Necessity of a functional octadecanoic pathway for indole alkaloid synthesis by *Catharanthus roseus* cell suspensions cultured in an auxin-starved medium. *Plant & Cell Physiology* 39: 220-225.

15 **Gish W, States DJ.** 1993. Identification of protein coding regions by database similarity search. *Nature (Genetics)* 3: 266-272.

20 **Goddijn OJM, de Kam RJ, Zanetti A, Schilperoort A, Hoge JHC.** 1992. Auxin rapidly down regulates transcription of the tryptophan decarboxylase gene from *Catharanthus roseus*. *Plant Molecular Biology* 18: 1113-1120.

Hashimoto T, Yamada Y. 1994. Alkaloid biogenesis: molecular aspects. *Annual Review of Plant Physiology and Plant Molecular Biology* 45, 257-285.

25 **Hashimoto T, Shoji T, Mihara T, Oguri H, Tamaki K, Suzuki K-i, Yamada Y.** 1998. Intraspecific variability of the tandem repeats in *Nicotiana* putrescine N-methyltransferases. *Plant Molecular Biology* 37: 25-37.

30 **Hibi N, Higashiguchi S, Hashimoto T, Yamada Y.** 1994. Gene expression in tobacco low-nicotine mutants. *Plant Cell* 6: 723-735.

Imanishi S, Hashizume K, Nakakita M, Kojima H, Matsubayashi Y, Hashimoto T, Sakagami Y, Yamada Y, Nakamura K. 1998a. Differential induction by methyl jasmonate of genes encoding ornithine decarboxylase and other enzymes involved in nicotine biosynthesis in tobacco cell cultures.

Plant Molecular Biology 38: 1101-1111.

5 Imanishi S, Hashizume K, Kojima H, Ichihara A, Nakamura K. 1998b. An mRNA of tobacco cell, which is rapidly inducible by methyl jasmonate in the presence of cycloheximide, codes for a putative glycosyltransferase. *Plant & Cell Physiology* 39: 202-211.

Junghans H, Metzlaff M. 1990. A simple and rapid method for preparation of total plant DNA. *Biotechniques* 8: 176.

10 Kanegae T, Kajiya H, Amano Y, Hashimoto T, Yamada Y. 1994. Species-dependent expression of the hyoscyamine 6 β -hydroxylase gene in the pericycle. *Plant Physiology* 105: 483-490.

Kutchan TM. 1995. Alkaloid biosynthesis - the basis for metabolic engineering of medicinal plants. *Plant Cell* 7, 1059-1070.

15 Kutchan TM. 1998. Molecular genetics of plant alkaloid biosynthesis. In: Cordell GA, ed., The Alkaloids, Chemistry and Biology. San Diego: Academic Press, 295-304.

20 Legg PD, Collins GB. 1971. Inheritance of percent total alkaloid in *Nicotiana tabacum* L. II Genetic effect of two loci in Burley 21 X LA Burley 21 populations. *Canadian Journal of Genetics and Cytology* 13: 287-291.

Legg PD, Chaplin JF, Collins GB. 1969. Inheritance of percent total alkaloids in *Nicotiana tabacum* L. *Journal of Heredity* 60: 213-217.

25 Lopes Cardosa MI, Meijer AH, Rueb S, Queiroz Machado J, Memelink J, Hoge JHC. 1997. A promoter region that controls basal and elicitor-inducible expression levels of NADPH: cytochrome P450 reductase (*Cpr*) from *Catharanthus roseus* binds nuclear factor GT-1. *Molecular & General Genetics* 25: 674-681.

30 Malik V, Watson MB, Malmberg RL. 1996. A tobacco ornithine decarboxylase partial cDNA clone. *Journal of Plant Biochemistry & Biotechnology* 5:109-112.

Malmberg RL, Watson MB, Galloway GL, Yu W. 1998. Molecular genetic analysis of plant

polyamines. *Critical Reviews in Plant Sciences* 17: 199-224.

Michael AJ, Furze JM, Rhodes MJC, Burtin D. 1996. Molecular cloning and functional identification of a plant ornithine decarboxylase. *Biochemical Journal* 314: 241-248.

5

Mizusaki S, Tanabe Y, Noguchi M, Tamaki E. 1973. Changes in the activities of ornithine decarboxylase, putrescine N-methyltransferase and N-methyl-putrescine oxidase in tobacco roots in relation to nicotine biosynthesis. *Plant & Cell Physiology* 14: 103-110.

10

Nam K.H. , S.H., Lee and J.H. Lee, A cDNA encoding arginine decarboxylase (GenBank U35367) from soybean hypocotyls. *Plant Physiol.*, 110: (1997) 714.

Nam K.H. , S.H. Lee and J.H. Lee, Differential expression of ADC mRNA during development and upon acid stress in soybean (*Glycine max*) hypocotyls. *Plant Cell Physiol.* 38 (1997) 1156-1166.

15

Ohnmeiss TE, McCloud ES, Lynds GY, Baldwin IT. 1997. Within-plant relationships among wounding, jasmonic acid, and nicotine: implications for defense in *Nicotiana sylvestris*. *New Phytologist* 137: 441-452.

20

Pasquali G, Goddijn OJM, de Waal A, Verpoorte R, Schilperoort RA, Hoge JHC, Memelink J. 1992. Coordinated regulation of two indole alkaloid biosynthetic genes from *Catharanthus roseus* by auxin and elicitors. *Plant Molecular Biology* 18: 1121-1131.

25

Pérez-Amador MA, Carbonell J. 1995. Arginine decarboxylase and putrescine oxidase in *Pisum sativum* L. Changes during ovary senescence and early stages of fruit development. *Plant Physiology* 107: 865-872.

30

Pérez-Amador MA, Carbonell J, Granell A. 1995. Expression of arginine decarboxylase is induced during early fruit development and in young tissues of *Pisum sativum* L. *Plant Molecular Biology* 28: 997-1009.

Primikirios, N.I. and K.A. Roubelakis-Angelakis. 1999. Cloning and expression of an arginine decarboxylase cDNA from *Vitis vinifera* L. cell-suspension cultures. *Planta* 208:574-582.

Riechers DE, Timko MP. 1999. Structure and expression of the gene family encoding putrescine N-methyltransferase in *Nicotiana tabacum*: new clues to the evolutionary origin of cultivated tobacco. *Plant Molecular Biology* 41: 387-401.

5 Rostogi R., J. Dulson and S.J. Rothstein, Cloning of tomato (*Lycopersicon esculentum* Mill.) arginine decarboxylase gene and its expression during fruit ripening. *Plant Physiol.*, 103 (1993) 829-834.

Saito K, Murakoshi I. 1998 Genes in alkaloid metabolism. In: Roberts MF, Wink M, eds. 10 Alkaloids: Biochemistry, Ecology, and Medicinal Applications. New York: Plenum Press, 147-157.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

15 Soyka S and A.G. Heyer. 1999. *Arabidopsis* knockout mutation of *ADC2* gene reveals inducibility by osmotic stress. *FEBS Lett* 458: 219-223.

Stim K.P. and G.N. Bennett, Nucleotide sequence of the *adi* gene, which encoded the biodegradative acid-induced arginine decarboxylase of *Escherichia coli*. *J. Bact.*, 175 (1993) 1221-20 1234.

Suzuki K, Yamada Y, Hashimoto T. 1999. Expression of *Atropa belladonna* putrescine N-methyltransferase gene in root pericycle. *Plant & Cell Physiology* 40: 289-297.

25 Wang J. 1999. Characterization of a cDNA (NtcADC1) and two nuclear genes (NtgADC1 and NtgADC2) encoding arginine decarboxylase, a key enzyme in alkaloid and polyamine biosynthesis in tobacco (*Nicotiana tabacum* L.). *M.S. Thesis, University of Virginia, Charlottesville, VA*.

Wang J, Sheehan M, Brookman H, and Timko MP. 2000. Characterization of cDNAs 30 Differentially Expressed in Roots of Tobacco (*Nicotiana tabacum* cv Burley 21) During the Early Stages of Alkaloid Biosynthesis. *Plant Science In press*

Watson M.B. and R.L. Malmberg, Regulation of *Arabidopsis thaliana* (L.) Heynh arginine decarboxylase by potassium deficiency stress. *Plant Physiol.*, 111 (1996) 1077-1083.

Watson M.B., W. Yu, G. Galloway and R.L. Malmberg, Isolation and characterization of a second arginine decarboxylase cDNA from *Arabidopsis* (Accession No. AF009647 (PGR97-114). Plant Physiol., 114 (1997) 1569.

5

Watson M.B., K. K. Emory, R.M. Piatak and R.L. Malmberg, 1998. Arginine decarboxylase (polyamine synthesis) mutants of *Arabidopsis thaliana* exhibit altered root growth. Plant. J. 13: 231-239.

10 Waterman PM. 1998. Chemical taxonomy of alkaloids. In: Roberts MF, Wink M, eds., Alkaloids: Biochemistry, Ecology, and Medicinal Applications. New York: Plenum Press, 87-107.

15 Cizewski-Culotta, V. and Hamer, D.H. 1989. Fine mapping of a mouse metallothionein gene metal response element. Mol. Cell. Biol. 9: 1376-1380.

Dellaporta, S.L., Wood, J. and Hicks, J.B. 1983. A plant DNA minipreparation: version II. Plant Mol. Biol. Rep. 1: 19-21.

20 Foster, R., Izawa, T. and Chua, N.H. 1994. Plant bZIP proteins gather at ACGT elements. FASEB J. 8: 192-200.

Gerstel, D.U. 1960. Segregation in new allopolyploids of *Nicotiana*. I. Comparison of 6 x (*N. tabacum* x *tomentosiformis*) and 6 x (*N. tabacum* x *otophora*). Genetics 45: 1723-1734.

25

Gerstel, D.U. 1963. Segregation in new allopolyploids of *Nicotiana*. II. Discordant ratios from individual loci in 6 x (*N. tabacum* x *N. sylvestris*). Genetics 48: 677-689.

30 Hashimoto, T., Tamaki, K., Suzuki, K. and Yamada, Y. 1998b. Molecular cloning of plant spermidine synthases. Plant Cell Physiol. 39: 73-79.

Hibi, N., Fujita, T., Hatano, M., Hashimoto, T. and Yamada, Y. 1992. Putrescine - methyltransferase in cultured roots of *Hyoscyamus albus*. n-Butylamine as a potent inhibitor of the transferase both *in vitro* and *in vivo*. Plant Physiol. 100: 826-835.

Kenton, A., Parok nny, A.S., Gleba, Y.Y. and Bennett, M.D. 1993. Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. Mol. Gen. Genet. 240: 159-169.

5 Kim, S.-R., Choi, J.-L., Costa, M.A. and An, G. 1992. Identification of G-box sequence as an essential element for methyl jasmonate response of potato proteinase inhibitor II promoter. Plant Physiol. 99: 627-631.

Kinoshita, T., Imamura, J., Nagai, H. and Shimotohno, K. 1992. Quantification of gene expression over a wide range by the polymerase chain reaction. Anal. Biochem. 206: 231-235.

10 Lalanne, E., Mathieu, C., Vedel, F. and De Paepe, R. 1998. Tissue-specific expression of genes encoding isoforms of the mitochondrial ATPase β subunit in *Nicotiana sylvestris*. Plant Mol. Biol. 38: 885-888.

15 Legg, P.D. and Collins, G.B. 1971. Inheritance of percent total alkaloids in *Nicotiana tabacum* L. II. Genetic effect of two loci in Burley 21 x LA Burley 21 populations. Can. J. Genet. Cytol. 13: 287-291.

Leitch, I.J. and Bennett, M.D. 1997. Polyploidy in angiosperms. Trends Plant Sci. 2: 470-476.

20 Li, W.-H. and Graur, D. 1991. Fundamentals of Molecular Evolution. Sinauer Associates, Inc., Sunderland, Massachusetts.

Mason, H.S., DeWald, D.B. and Mullet, J.E. 1993. Identification of a methyl jasmonate-25 responsive domain in the soybean *vspB* promoter. Plant Cell 5: 241-251.

Menkens, A.E., Schindler, U. and Cashmore, A.R. 1995. The G-box: a ubiquitous regulatory element in plants bound by the GBF family of bZip proteins. Trends Biochem. Sci. 20: 506-510.

30 Nakajima, K., Hashimoto, T. and Yamada, Y. 1993. Two tropinone reductases with different stereospecificities are short-chain dehydrogenases evolved from a common ancestor. Proc. Natl. Acad. Sci. USA 90: 9591-9595.

Okamura, J.K. and Goldberg, R.B. 1985. Tobacco single-copy DNA is highly homologous to

sequences present in the genomes of its diploid progenitors. *Mol. Gen. Genet.* 198: 290-298.

Ramsey, J. and Schemske, D.W. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annu. Rev. Ecol. Syst.* 29: 467-501.

5

Rouster, J., Leah, R., Mundy, J. and Cameron-Mills, V. 1997. Identification of a methyl jasmonate-responsive region in the promoter of a lipoxygenase 1 gene expressed in barley grain. *Plant J.* 11: 513-523.

10 Saunders, J.W. and Bush, L.P. 1979. Nicotine biosynthetic enzyme activities in *Nicotiana tabacum* L. genotypes with different alkaloid levels. *Plant Physiol.* 64: 236-240.

15 Shinshi, H., Wenzler, H., Neuhaus, J.-M., Felix, G., Hofsteenge, J. and Meins, F. 1988. Evidence for N- and C-terminal processing of a plant defense-related enzyme: Primary structure of tobacco prepro- β -1,3-glucanase. *Proc. Natl. Acad. Sci. USA* 85: 5541-5545.

Sperisen, C., Ryals, J. and Meins, F. 1991. Comparison of cloned genes provides evidence for intergenic exchange of DNA in the evolution of a tobacco glucan endo-1,3- β -glucosidase gene family. *Proc. Natl. Acad. Sci. USA* 88: 1820-1824.

20

Staiger, D., Kaulen, H. and Schell, J. 1989. A CACGTG motif of the *Antirrhinum majus* chalcone synthase promoter is recognized by an evolutionary conserved nuclear protein. *Proc. Natl. Acad. Sci. USA* 86: 6930-6934.

25

Suzuki, K., Yun, D.-J., Chen, X.-Y., Yamada, Y. and Hashimoto, T. 1999b. An *Atropa belladonna* hyoscyamine 6 β -hydroxylase gene is differentially expressed in the root pericycle and anthers. *Plant Mol. Biol.* 40: 141-152.

Thiele, D.J. 1992. Metal-regulated transcription in eukaryotes. *Nucl. Acids Res.* 20: 1183-1191.

30

Thompson, J.D. and Lumaret, R. 1992. The evolutionary dynamics of polyploid plants: origins, establishment and persistence. *Trends Ecol. Evol.* 7: 302-307.

Vaucheret, H., Vincentz, M., Kronenberger, J., Caboche, M. and Rouze, P. 1989. Molecular

cloning and characterisation of the two homeologous genes coding for nitrate reductase in tobacco.

Mol. Gen. Genet. 216: 10-15.

Williams, M.E., Foster, R. and Chua, N.H. 1992. Sequences flanking the hexameric G-box core

5 CACGTG affect the specificity of protein binding. Plant Cell 4: 485-496.

What is claimed is:

1. An isolated DNA molecule comprising the nucleotide sequence of (SEQ. ID. NO. 2), (SEQ. ID. NO. 5), (SEQ. ID. NO. 8), (SEQ. ID. NO. 11), (SEQ. ID. NO. 13), (SEQ. ID. NO. 15), (SEQ. ID. NO. 18), (SEQ. ID. NO. 21), (SEQ. ID. NO. 23), (SEQ. ID. NO. 25) or (SEQ. ID. NO. 26) or comprising a nucleotide sequence encoding the amino acid sequence encoded by (SEQ ID NO. 3), (SEQ. ID. NO. 6), (SEQ ID. NO. 9), (SEQ. ID. NO. 12), (SEQ. ID. NO. 14), (SEQ. ID. NO. 16), (SEQ. ID. NO. 19), (SEQ. ID. NO. 22) OR (SEQ. ID. NO. 24).
- 10 2. A vector comprising the isolated DNA molecule of claim 1 operably linked to sequences capable of directing the transcription of a mRNA encoded by said isolated DNA molecule.
- 15 3. An isolated DNA molecule comprising a DNA sequence complementary to the nucleotide sequence of claim 1.
4. A vector comprising the isolated DNA molecule of claim 3 operably linked to sequences capable of directing the transcription of a mRNA encoded by said isolated DNA molecule.
5. A cultured transgenic tobacco cell stably transformed with the vector of claim 2.
- 20 6. A cultured transgenic tobacco cell stably transformed with the vector of claim 4.
7. A transgenic tobacco plant stably transformed with the vector of claim 2.
- 25 8. A transgenic tobacco plant stably transformed with the vector of claim 4.
9. The isolated DNA molecule of claim 1, wherein the isolated DNA molecule comprises the nucleotide sequence of (SEQ ID NO:).
- 30 10. A vector comprising the isolated DNA molecule of claim 9 operably linked to sequences capable of directing the transcription of a mRNA encoded by said isolated DNA molecule.
11. An isolated DNA molecule comprising a DNA sequence complementary to the nucleotide sequence of the isolated DNA molecule of claim 9.

12. An isolated DNA sequence comprising about a fifteen to about a twenty-five base pair oligonucleotide sequence identical to any consecutive about fifteen to about twenty-five base pair sequence found in (SEQ. ID. NO. 2), (SEQ. ID. NO. 5), (SEQ. ID. NO. 8), (SEQ. ID. NO. 11), (SEQ. ID. NO. 13), (SEQ. ID. NO. 15), (SEQ. ID. NO. 18), (SEQ. ID. NO. 21), (SEQ. ID. NO. 23),
5 (SEQ. ID. NO. 25) or (SEQ. ID. NO. 26).

13. A cultured transgenic tobacco cell stably transformed with the vector of claim 10.

14. A transgenic tobacco plant stably transformed with the vector of claim 10.

10 15. A vector comprising a DNA sequence which encodes an antisense mRNA which is complementary to a fragment of a mRNA encoded by the isolated DNA molecule of claim 1, wherein said sequence is operably linked to sequences capable of directing the transcription of said antisense mRNA in tobacco cells and wherein the expression of said antisense mRNA in tobacco
15 cells is sufficient to provide for reduced nicotine content in tobacco cells which are stably transformed with said vector as compared to untransformed tobacco cells.

16. A cultured transgenic tobacco cell stably transformed with the vector of claim 15.

20 17. An isolated and purified protein comprising the amino acid sequence identified in (SEQ ID NO. 3), (SEQ. ID. NO. 6), (SEQ ID. NO. 9), (SEQ. ID. NO. 12), (SEQ. ID. NO. 14), (SEQ. ID. NO. 16), (SEQ. ID. NO. 19), (SEQ. ID. NO. 22) or (SEQ. ID. NO. 24).

25 18. A method for regulating gene expression in a plant comprising functionally linking an alkaloid gene promoter to a nucleic acid encoding a protein, wherein the promoter comprises a nucleic acid sequence selected from the group consisting of the sequences identified in (SEQ ID NO. 1), (SEQ. ID. NO. 4), (SEQ ID. NO. 7), (SEQ. ID. NO. 10), (SEQ. ID. NO. 17), and (SEQ. ID. NO. 20).

30 19. The method of claim 18, wherein the nucleic acid encoding a protein encodes a protein involved in the biosynthesis of alkaloids in plants.

20. A plant transformed by the method of claim 18.

1 / 21

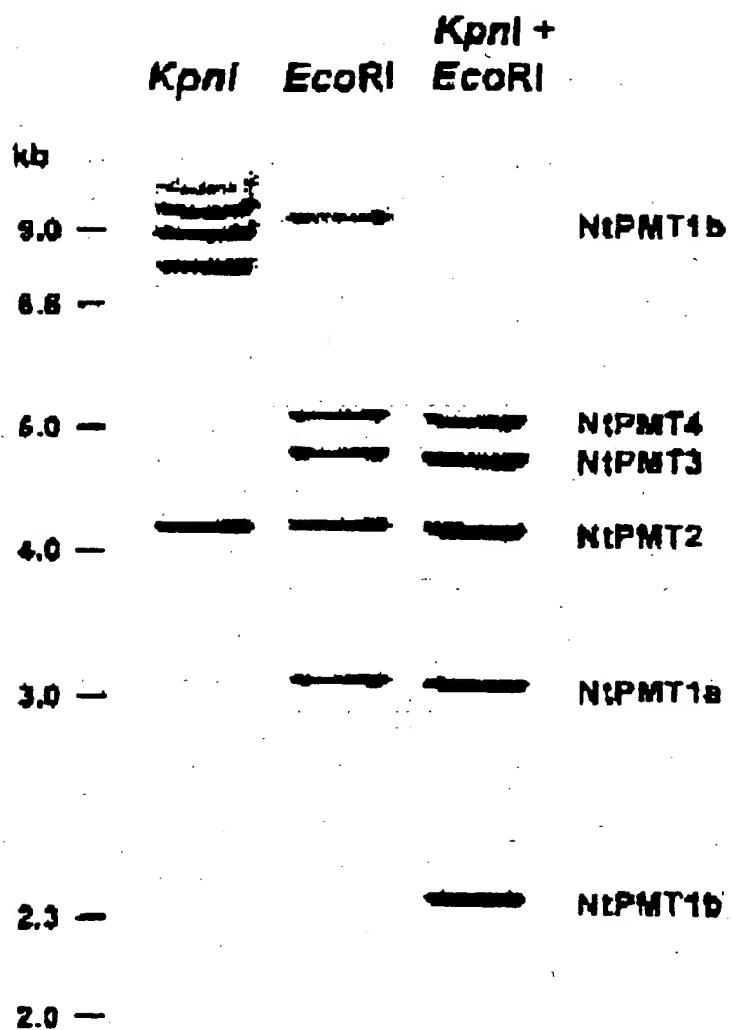


Figure 1

		Exon 1		
NIPMT4	M E V I S T N T N G S T I F K N G A I P M N G H Q S G T I S K H L N G Y Q N G T S K H Q			
NIPMT3	M E V I S T N T N G S T I F K N G A I P M			
NIPMT1a	M E V I S T N T N G S T I F K N G A I P M N G H Q S G T S			
NIA411	M E V I S T N T N G S T I F K N G A I P M N G H Q N G T S			
NIPMT2	M E V I S T N T N G S T I F K N G A I P M N G H Q N G T S			
NIPMT4	G	N G H Q N G	E H O N G H O N G T S E	
NIPMT3	G	N G H Q N G	E H O N G H O N G T S E	
NIPMT1a	N G H Q N G	E H R N G H O N G T S E	
NIA411	N G H Q N G	E H R N G H O N G T S E	
NIPMT2	H Q N G H	H Q N G H	
NIPMT4	Q Q N G T I S H D N G N E	L L G N S N S I K P G W F S E F S A L W P G		
NIPMT3	Q Q N G T I S H D N G N E	L L G S S S N S I K P G W F S E F S A L W P G		
NIPMT1a	Q Q N G T I S H D N G N E	L L G S S S I K P G W F S E F S A L W P G		
NIA411	Q Q N G T I S H D N G N E	L L G S S S I K P G W F S E F S A L W P G		
NIPMT2	Q Q N G T I S H D N G N E	L L G R S N S I K P G W F S E F S A L W P G		
		Exon 2	Exon 3	
NIPMT4	E A F I S L K V E K L L F Q G K S D Y Q D V M L F E S A T Y G K V I L T L D G A I Q H T E N G G F P			
NIPMT3	E A F I S L K V E K L L F Q G K S D Y Q D V M L F E S A T Y G K V I L T L D G A I Q H T E N G G F P			
NIPMT1a	E A F I S L K V E K L L F Q G K S D Y Q D V M L F E S A T Y G K V I L T L D G A I Q H T E N G G F P			
NIA411	E A F I S L K V E K L L F Q G K S D Y Q D V M L F E S A T Y G K V I L T L D G A I Q H T E N G G F P			
NIPMT2	E A F I S L K V E K L L F Q G K S D Y Q D V M L F E S A T Y G K V I L T L D G A I Q H T E N G G F P			
NIPMT4	Y T I E M I V H L I P L G S I P N P I K K V L I I G G G I G F T L F E M I L R Y P T I I E K I I D I V E I D			
NIPMT3	Y T I E M I V H L I P L G S I P N P K K V L I I G G G I G F T L F E M I L R Y P T I I E K I I D I V E I D			
NIPMT1a	Y T I E M I V H L I P L G S I P N P K K V L I I G G G I G F T L F E M I L R Y P T I I E K I I D I V E I D			
NIA411	Y T I E M I V H L I P L G S I P N P K K V L I I G G G I G F T L F E M I L R Y P T I I E K I I D I V E I D			
NIPMT2	Y T I E M I V H L I P L G S I P N P K K V L I I G G G I G F T L F E M I L R Y P T I I E K I I D I V E I D			
		Exon 4	Exon 5	
NIPMT4	D V V V D V S R K S F P Y L A A N F N D P R V T L L V L G D G A A F V K A A Q A G Y Y D A I I V D			
NIPMT3	D V V V D V S R K F F P Y L A A N F N D P R V T L L V L G D G A A F V K A A C A G Y Y D A I I V D			
NIPMT1a	D V V V D V S R K F F P Y L A A N F N D P R V T L L V L G D G A A F V K A A Q A G Y Y D A I I V D			
NIA411	D V V V D V S R K F F P Y L A A N F N D P R V T L L V L G D G A A F V K A A Q A G Y Y D A I I V D			
NIPMT2	D V V V D V S R K F F P Y L A A N F N D P R V T L L V L G D G A A F V K A A Q A G Y Y D A I I V D			
NIPMT4	S S D P I I G P A K D L F E R P F F E A V A K A L R P G G V V C T Q A E S I I W L H M H I I K Q I I			
NIPMT3	S S D P I I G P A K D L F E R P F F E A V A K A L R P G G V V C T Q A E S I I W L H M H I I K Q I I			
NIPMT1a	S S D P I I G P A K D L F E R P F F E A V A K A L R P G G V V C T Q A E S I I W L H M H I I K Q I I			
NIA411	S S D P I I G P A K D L F E R P F F E A V A K A L R P G G V V C T Q A E S I I W L H M H I I K Q I I			
NIPMT2	S S D P I I G P A K D L F E R P F F E A V A K A L R P G G V V C T Q A E S I I W L H M H I I K Q I I			
		Exon 6		
NIPMT4	A N C I R Q V F K G S V N Y A W T T V P T Y P T G V I I G Y M L C S T E G P E I V D F K N P V N P I I D			
NIPMT3	A N C I R Q V F K G S V N Y A W T T V P T Y P T G V I I G Y M L C S T E G P E I V D F K N P V N P I I D			
NIPMT1a	A N C I R Q V F K G S V N Y A W T T V P T Y P T G V I I G Y M L C S T E G P E I V D F K N P V N P I I D			
NIA411	A N C I R Q V F K G S V N Y A W T T V P T Y P T G V I I G Y M L C S T E G P E I V D F K N P V N P I I D			
NIPMT2	A N C I R Q V F K G S V N Y A W T T V P T Y P T G V I I G Y M L C S T E G P E I V D F K N P V N P I I D			
		Exon 7		
NIPMT4	K E T T I Q V K S K L A P L K F Y N S D I H K A A F I I L P S F A R S M I E S			
NIPMT3	K E T T I Q V K S K L A P L K F Y N S D I H K A A F I I L P S F A R S M I E S			
NIPMT1a	K E T T I Q V K S K L A P L K F Y N S D I H K A A F I I L P S F A R S M I E S			
NIA411	K E T T I Q V K S K L A P L K F Y N S D I H K A A F I I L P S F A R S M I E S			
NIPMT2	K E T T I Q V K S K L A P L K F Y N S D I H K A A F I I L P S F A R S M I E S			
		Exon 8		

Figure 2

3 / 21

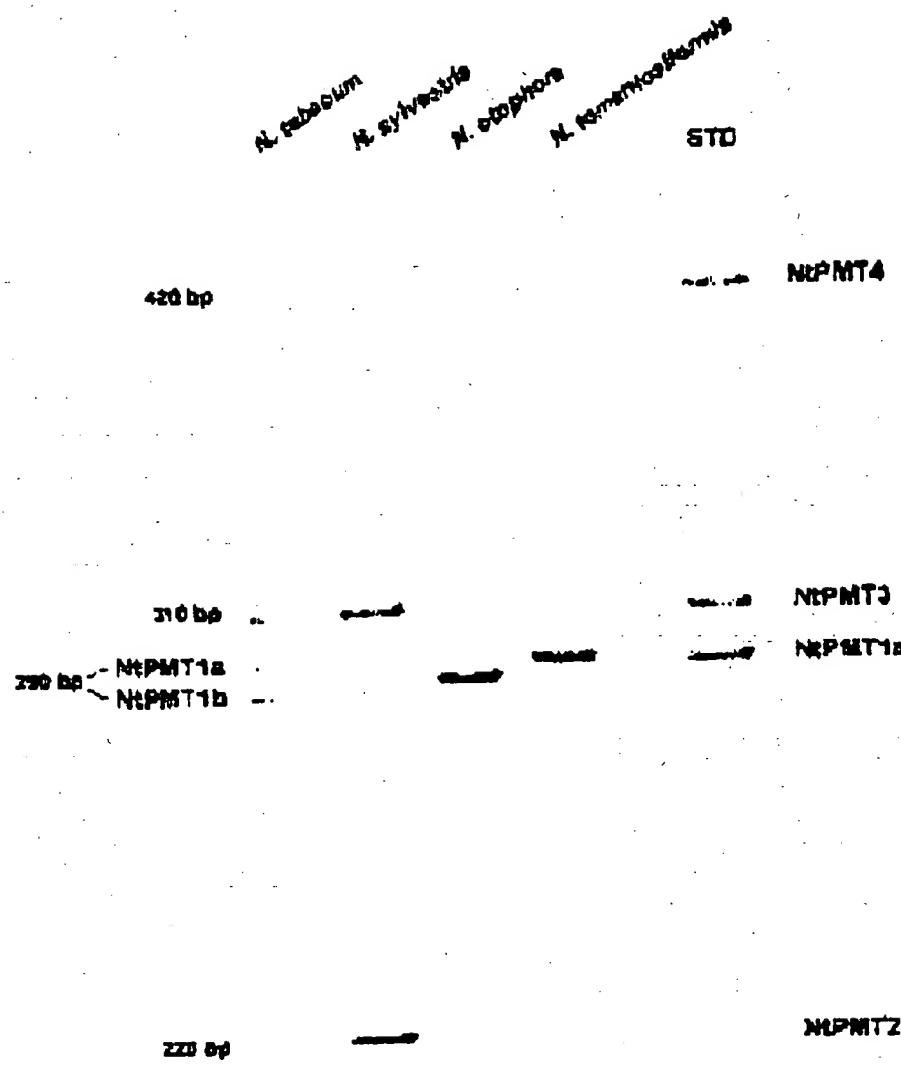


Figure 3

4 / .21

Figure 4

5 / 21

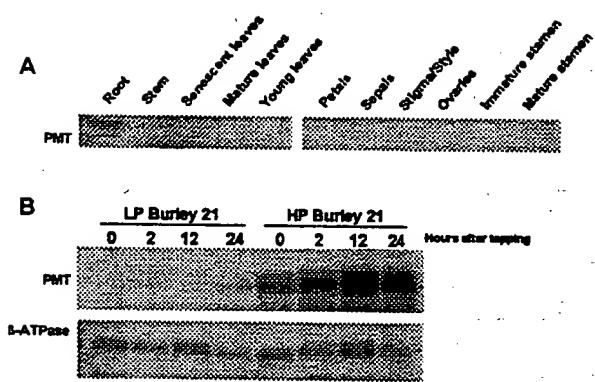


Figure 5

6 / 21

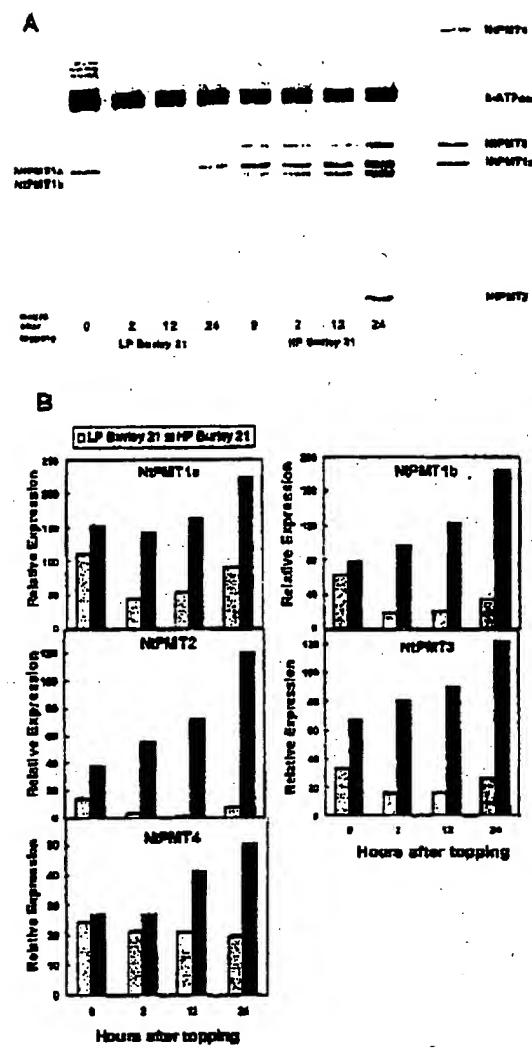


Figure 6

7 / 21

NtADC1 319
 ttcacgttctttcaattccataaaagaaaaacccttcgttag
 gtttccgtcttatTTT--cttttttacgcttc 78

NtADC2 c.....
 tc...a.....c
 ..t... 80

NtADC1
 ctcttctgatataatctgtatgggtttttcttg
 ttcgaattttagattgtttgccttaataacctgta
 acctta 158

NtADC2 a.....
 t.....a..
 160

NtADC1
 taattctctgtttaaacccaaaaacttagcttctcg
 aagtcaagggtgggatattggatcgtgtaaagagtgt
 gttaga 238

NtADC2 -

 t.....
 239

NtADC1
 aggtgattatctttgattcagttcccttttgccttc
 ttttgggggttagccggggcctcggcctcggcgggt
 tttaat 318

NtADC2 g.....

NtADC1 399
 agccccatctattacaaccattggcaaaaacatca
 ttaaatctgtacaaagcaaaccctaatttagttaa
 ttttct 398

NtADC2 t.....
 a.....
 399

1
 M P A L G C C V D A T -
 V S P P

NtADC1
 gtatttttgcattttaaacagaagaagagAGTC
 CGGCCCTAGGTTGGCGTAGACGCTACT---
 GTTTCCCCTCC 475

NtADC2 a.....t.....
 T.....T....G..GTT.....
 479

1
 M * * * * * * * * A V
 * * * *

16 L G Y A F S R D S S
 L P A P E F F T S G V P P
 T N S A

NtADC1
 CTCGGCTATGCCTCTCTCGGGATAGCTCTCTCCCG

Figure 7 (a)

8 / 21

Figure 7 (b)

9 / 21

Figure 7(c)

Figure 7 (d)

363	Q E Y A S T V V Q A	NtADC1
V Q Y V C D R K G V K H P	TTCTTCACATCTGCTTCTGGTGGCCTCCAATCCATG	
V I C	GCGGAGACGCTCAATGAAGATGCCCTTGCTGATTACC	
NtADC1	GCAATT 1755	
TTCAAGAATAACGCCTCCACAGTTGTCCAGGCAGGTCA	NtADC2	
ATATGTTGCGACCGTAAGGGCGTGAAGCACCCAGTG	
ATTTGC 1595 C	
NtADC2 1759	
..... T G T	417 * * * * * * * * * *	
..... A T A	* * * * * * * * * * * * * * *	
..C... 1599	* * * *	
364	*	443 S A A A V R G E Y E
*	*	T C V L Y S D Q L K Q R C
*	*	V D Q
389	S E S G R A I V S H	NtADC1
H S I L I F E A V S A S S	TATCTGCTGCTGCAGTTCTGGAGAGTACGGAGACGTG	
H S C S	TGTACTTTACTCTGATCAGTTGAAACAGAGATGTGTG	
NtADC1	GATCAG 1835	
AGCGAAAGTGGCAGGGCAATTGTTCTCATCACTCAA	NtADC2	
TTCTGATTTCGAACGCCGTGTCTGCTTAGTCACTC T A	
ATGTTTC 1675	
NtADC2 1839	
.....	444 * * * * * * * * * *	
.....	* * * * * * * * * * * * * *	
..... 1679	* * * *	
390	*	469 F K E G S L G I E H
*	*	L A A V D S I C D F V S K
*	*	A M G A
416	S S H L S S G G L Q	NtADC1
S M A E T L N E D A L A D	TTTAAAGAAGGGTCCCTGGGTATTGAACATCTGCTG	
Y R N L		

Figure 7 (e)

NtADC2

T.....C.....

G.....A.....

2079

524 * * * * * * * * * *

* * * * * * * * * *

* * * * *

549 G K V D K F I G G E

S S L Q L H E L G S N G D

G G G Y

NtADC1

GGGAAGGTTGATAAGTTCATGGTGGCGAACATCAAGCT

TGCAGCTGCATGAATTGGGAAGTAATGGCGATGGTGG

TGGGTA 2155

NtADC2

C..A.....

T.. 2159

550 * * * * * * * * * *

* * * P * * * * * * * *

* * * *

576 Y L G M F L G G A Y

E E A L G G L H N L F G G

P S V V

NtADC1

TTATCTGGGGATGTTTTGGGTGGGGCTTATGAGGAG

GCGCTCGGAGGACTCCACAACCTGTTGGTGGACCAA

GCGTGG 2235

NtADC2

Figure 7 (f)

Figure 7(g)

14 / 21

709 T G E D E I W S Y C
T A ***

NtADC1
ACAGGGGAGGGATGAGATTGGTCTATTGCACTGCTT
GAagtgttgcgttagcatctccagtttagttgtcg
tcgaag 2635
NtADC2
.....T
GA.....c.....
....g. 2639
710 * * * * * * * *
* * * * ***

720
NtADC1
ttgtctgttttgaataataacccttagttggatgt
ttttct
2678
NtADC2
.....aataaata.....
.....
2682

Figure 7(h)

Figure 8

16 / 21

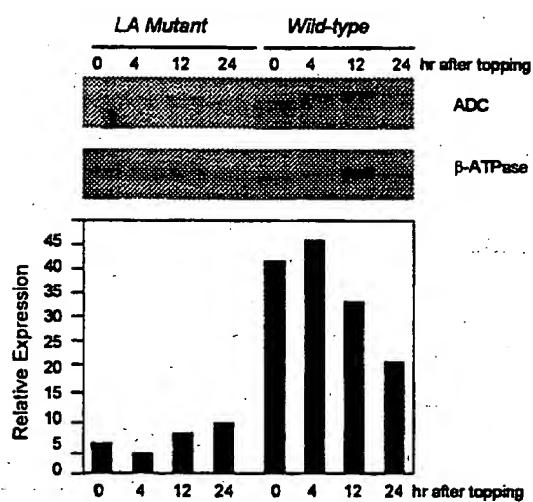


Figure 9

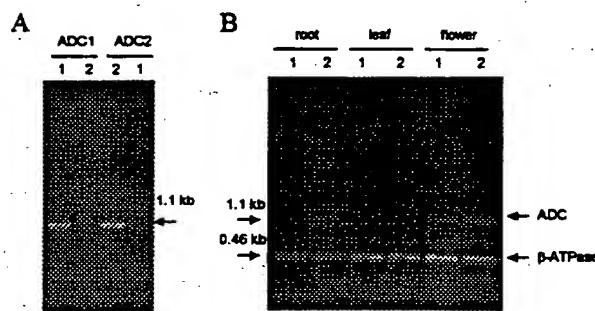


Figure 10

17 / 21

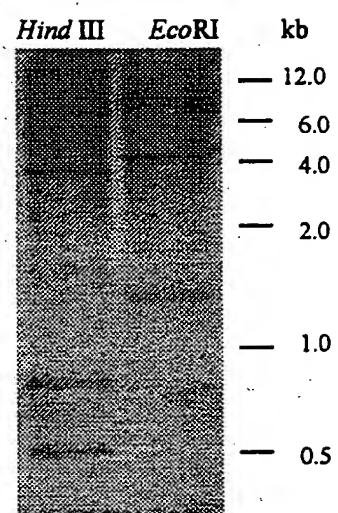


Figure 11

18 / 21

-135
 ODC2 CCTTACCCCTT CAACAGCTAT TTCTCTAAA AAAAAAAA AGAAAGAAA TACTAGCTG ATTACACAT ATTATCACTA GTAGTATCAC TTTCCTCCCC
 ODC1 AAAGTACTT TAACCATA TGACCGTGTG AAAAGCCCA AAAAACAGTT TTTTTTATT TTTTTTATT TCCCTCCAAA AACACATTAA AGGTATTTT

-35 TATA box +1+ * * *
 ODC2 TCTCTCTCT GATAAACATT TTAGAGCTT TCCCTCTC AAAGGGAAAC AGAGGAACAT TCATATTCAT GAATCCCTG TTCTCTCTT TTCCCTTTGA
 ODC1 TAAGCACATT TTCTCTCTT CCTTCCCG CGGTATTCG ATCCCTTC CAATGCCTC ACCATGCTG GCACCCCTT CTGCTTCTT TTCCCTTTGA

65
 ODC2 TTCCCTCTC TCATTTACTC TTCTCTCTC TTCTCTCTT TTGGATGGCCG GCACAAACAT CATCCTTCC GGTTGAAAC CGGGGGCCAT TTTCAGTC
 ODC1 TTCCCTCTC TCATTTACTC CTATCTTC TTACTCTCTT TTGGATGGCCG GCACAAACAT CATCCTTCC GGTTGAAAC CGGGGGCCAT TTTCAGTC
 pODC2 K A G Q T I I V S G L N P A A I L O S
 pODC1 C

165
 ODC2 TTCCCTCTC TCATTTACTC TTCTCTCTC TTCTCTCTT TTGGATGGCCG GCACAAACAT CATCCTTCC GGTTGAAAC CGGGGGCCAT TTTCAGTC
 ODC1 TTCCCTCTC TCATTTACTC CTATCTTC TTACTCTCTT TTGGATGGCCG GCACAAACAT CATCCTTCC GGTTGAAAC CGGGGGCCAT TTTCAGTC
 pODC2 K A G Q T I I V S G L N P A A I L O S
 pODC1 C

265
 ODC2 ACATTTGCCG GGGGAGCTTC TTCTACAGCG CGGGGGCGG CGGGAAACGG CACCAAGAAA GTCATCTTC TTCTCAAGAGA TTCTCTACAA GATTTCATGT
 ODC1 ACATTTGCCG GGGGAGCTTC TTCTACAGCG CGGGGGCGG CGGGAAACGG CACCAAGAAA GTCATCTTC TTCTCAAGAGA TTCTCTACAA GATTTCATGT
 pODC2 T I G G G A S P T A A A A E N G T R K V I P L S R D A L Q D E M
 pODC1 D

365
 ODC2 TATCAAATCAT AAACCAAAA TTACAAGATG AGAACAAACCC TTCTTACGTG CTAGACTTGG GTGAGGTTTG TTCTCTTATG GACCAATGGA AATTTGCTT
 ODC1 TATCAAATCAT AAACCAAAA TTACAAGATG AGAACAAACCC TTCTTACGTG CTAGACTTGG GTGAGGTTTG TTCTCTTATG GACCAATGGA AATTTGCTT
 pODC2 L S I I T Q K L Q D E K Q P F T V L D D L G E V V S L H D Q W K S A L
 pODC1 Y

465
 ODC2 CCCAAATATC CTGCTTAAATC ACCTGTTAAAT ATGTAACCCCT GAACCGCTGTG TTCTCTCAAT TTATCTGCT ATGGGCTCAA ATTTTGATTG TGCTAGCCGA
 ODC1 CCCAAATATC CTGCTTAAATC ACCTGTTAAAT ATGTAACCCCT GAACCGCTGTG TTCTCTCAAT TTATCTGCT ATGGGCTCAA ATTTTGATTG TGCTAGCCGA
 pODC2 P H I R P E Y A V K C N P E P S F L S I L S A M G S E D C A S R
 pODC1 S L

565
 ODC2 GCTGAAATTG ACTATGTTT ATCTCTTGG ATTCACCTG ACCCTATGTT TTGGAAAT CCATGAAAC CGGAATCCGA TATTATTTTT GCACCAAG
 ODC1 GCTGAAATTG ACTATGTTT ATCTCTTGG ATTAAGAAAAG GAGAGAGTC ATGGGTTAC TGAGTTTATG GAAAGTTGG GAAATTAAATA TTGGGTTGT
 pODC2 A X I E Y V L S L G I S P D R I V F A N P C K F S S D I X F A A K
 pODC1 N K K R E R S H M G Y L X .

665
 ODC2 TTGGGTGAA TTCTACACCC TTGATTCTG AAGAGGAGGT TTACAAGATC CGAAAGCATC ACCCGAAATC CGAACCTCTG CTGGCATCA ACCCCATGCT
 ODC1 TTGGGTGAA TTCTACACCC TTGATTCTG AAGAGGAGGT TTACAAGATC ACCCGAAATC CGAACCTCTG CTGGCATCA ACCCCATGCT
 pODC2 V G V N L T T Y D S E D V T K I R K H E P K S E L L L R I K P H L

765
 ODC2 CGACGGCAGC CGGAGATGCC CAATGGGCC GAAATACGGC CGCTTCCAG AAGAAATCGA CGCGCTGCTC CGGGCAGCTC AGCCCGCCCG TTCTACCGTA
 ODC1 ATGAAAGAGC CGAAGACAC CAAGACCACT GATTCCCCAA ACACCAAAATT TCATTTTTT TAACGTTTT CTTCCTGTTGTTGTTAA TTACGTTTT
 pODC2 D G H A R C P H G P K Y G A L P E E V D P L L R A A Q A A R L T V

865
 ODC2 TCGGGCGTGT CAPTOCACAT CGGTAGCGGA GATGCCGATT CAAACCTTA TCTGGGGCC ATAGCGGGG CTAAGGAAGT TTGGAAACA GTCGCTAAAC
 ODC1 CTGGCTCTT TTAGAAATT ATTITTTATT TATTATTTAA ATAGATTTAA CATACTTTTT TTACTCTAA ATAATATATG TCATTTTTTT ATTCGCTACT
 pODC2 S G V S F H I G S C D A D S M A T L G A I A A K E V F I T A A K

965
 ODC2 TCGGGATGTC GAAATGACT TTCTCTAGG TGGGGGGGG GTTACATCC GCCCACCGT TCACACCCG CGGGCTGCC GTAAATCG CTTAAAACA
 ODC1 CGGGCACGTCA CGAGGGAGTC CATTGACAC ACCTTGTAGA TTGGCTGCTAT TGGCTCAATT GGACCAAGT TCATGATACG TATGATGCTCA
 pODC2 L G H S E M T V L D V G G G F T S G H Q F T T A A V A V K S A L K Q

1065
 ODC2 ACATTCGAT GACGAACCGG AGTTGACANT CATACTGAA CGGGGTGGGT TTGGTGGAGA GACGGGGTTT ACCTTGGCAA CGGCGATTAT AGGGAAAGA
 ODC1 ATAGGAACCTC TTCTAACTTT AGGTGTCTAA ATGAAAGATC TGCCACCT TCAGTGTCTC CGTATGTTT CAGCCAAAT AAATGAAGCC AAATGATGTC
 pODC2 E F D D K P E L T I I A E P G R E F A E T A F T L A T T T I G K R

1165
 ODC2 GTGGGGGTG AATTTGGGAA GTATTTGGATT AACGACGGGC TGACGGTTC GATGAACTG GTACTTACG ACCATGGCAC CGTGAAATGCA ACCGGGTTAG
 ODC1 AATAAAGCGA TCTGCTAGA ACCACGGAC TCAGGGATG CTCTACACCT CTCCTCCGGT CAACAGAAATT CCTACTCGG AGTTGGTTT CGAAGACCAA
 pODC2 V R G E L R E Y V I H D G E Y G S H N C V L T D H A T V H A T T P L

1265
 ODC2 CTGCTCTGCA GAAATGACTG AACCTGACTT CGGGGGGGTC GAAACGTTT CGGAGGACTG TTGGGGGGC CACTTGTGAT GCTCTGATA CTGTTTAAAG
 ODC1 TAATAATAGA GTGAAACCTT CCTTGAATAA GGGATTCAA AAAAGGTTGA CTGGAAACAC CAGCIAAAAT TAATTCCTG TGCGACACT GTAAATTTAA
 pODC2 A V L S H R S H V T C G G S K T F P T T V F G P T C D A L D T V L R

1365
 ODC2 GGATTTACCG TTACGGGAGC TGCAAGTTAA TGATTTGGCTG GTTTTTCTA ATATGGTGC TTACTACTAA GCTGCTGGGT CCAATTTTAA TTGGATTTAAAT
 ODC1 TAATCCCTAT TTCAATTTC TCACCTTAAATG GGGGGGGGGCA CAACTCTAA CAACACCTAA TTCTTTGGAG GTGATTTAAAG GTGATGTCAC
 pODC2 D Y Q L P E L Q V H D W L V F P H N G A Y T K A A G S M F N G F N

1465
 ODC2 ACTTCGGCCA TTGTTACTCA CCTCGCTTAT TCCTTACCTAA CCTGATGAAAC CACCTGTTT AGGAATTACT ACCGTTGGTTT TGATGGTTTT TTCTCTCTT
 ODC1 AGCTCTAGCA ACTCTGGCTGG GGGCTTAAAT TAAGAACTTC AGCTTGGTTAT ATTGATTTT ATTGGCTTT TATCATGCTT TGATGTTTAT TTCTCTGGG
 pODC2 T S A I V T H L A T S Y P S .

1565
 ODC2 Poly A signal 1565
 ODC1 GGGATCTCTT TTCTTAAATT TGTTGTTTT GTGAGTAAATT TATATTCCTAA ATCAGCTGT AATCTCTTC TATGCGAAGGTTGCAAGG ATTTGCTAAT
 pODC2 AGCATATGT CTCTTGGCTC TTCTTAAATT CCTTAACTAA CTGCTGATAA ATTGTATCC TATCTCCGAC CCTCTCTGAGT CTTCCTGAG

1666
 ODC2 1666
 ODC1 TGTGATTTTC TCTAATATGG AATTTTTAA AATTAATTTA AGAAACATAA TGGGAAAGG GTTTGGGG TCATGATATT TGTTGACTA TAAAGGATC
 pODC2 GTAGTATGT TGCTCTTGC TACCAAGCATC ATAATATTC TTCTCTGAGA TAAACCCAGT TAGGCTACCA CCTTTGGT AAGGATTTAA TCACATATGT

Figure 12

19 / 21

1		50
N. tabacum cv Xanthi	[REDACTED]	GASPTA[A]G-[R]
N. tabacum cv BY2	[REDACTED]	GASPT-[A]G-[R]
N. tabacum cv SC58	[REDACTED]	GASPTA[A]G-[R]
D. stramonium	[REDACTED]	A---TP[P]W[D]H[Q][E]R
L. esculentum	[REDACTED]	A---PV[A]M[G]H[P]W[K]
S. cerevisiae	MSSTQVGNALSSSTTLVDL	INSTVTQKKYYKDETLHNLLLELK[NQD]
H. sapiens	MNF--GN-	
51		100
N. tabacum cv Xanthi	[REDACTED]	[REDACTED]-SND
N. tabacum cv BY2	[REDACTED]	[REDACTED]-SND
N. tabacum cv SC58	[REDACTED]	[REDACTED]-SND
D. stramonium	[REDACTED]	[REDACTED]-D
L. esculentum	[REDACTED]	[REDACTED]-D[E]
S. cerevisiae	LELEHEQAHPKIFQ	KARIGRINNETCDPGENSFIC[KR]FN
H. sapiens	EEFDCH--FLDEGFT	-KEILDOK[NEVSSS]D[DAG]A[DILKKHL
101		150
N. tabacum cv Xanthi	[REDACTED]	I-[REDACTED]
N. tabacum cv BY2	[REDACTED]	I-[REDACTED]
N. tabacum cv SC58	[REDACTED]	I-[REDACTED]
D. stramonium	[REDACTED]	[REDACTED]-M
L. esculentum	[REDACTED]	[REDACTED]-M
S. cerevisiae	NEVKE[R]KE	DTKV[L]AELEV[KV]DR[HMN]
H. sapiens	RALK[R]RT	DSKAIVKT[A]T[TG]K[Q]OLVQEV
151		200
N. tabacum cv Xanthi	[REDACTED]	V-[REDACTED]-Y-[REDACTED]-S
N. tabacum cv BY2	[REDACTED]	V-[REDACTED]-Y-[REDACTED]-S
N. tabacum cv SC58	[REDACTED]	V-[REDACTED]-Y-[REDACTED]-S
D. stramonium	[REDACTED]	[REDACTED]-Y-[REDACTED]-C
L. esculentum	[REDACTED]	[REDACTED]-Y-[REDACTED]-C
S. cerevisiae	VASHFERY[SKN]MKSF	UNVEELH[KI]FESQ
H. sapiens	P[E]S[I]QV[Q]KY[NN]QMM	F[RE]VLMVARAK[V]
201		250
N. tabacum cv Xanthi	[REDACTED]	D-[REDACTED]A-[REDACTED]
N. tabacum cv BY2	[REDACTED]	D-[REDACTED]A-[REDACTED]
N. tabacum cv SC58	[REDACTED]	D-[REDACTED]A-[REDACTED]
D. stramonium	[REDACTED]	[REDACTED]-T
L. esculentum	[REDACTED]	[REDACTED]-T
S. cerevisiae	-ATD[STO]RLST[CE]MNDV	KAIKELGENLA[TV]SA
H. sapiens	-ATD[SKV]RLSV[TL]RTSRL	[ER]KELNID[IV]C
251		300
N. tabacum cv Xanthi	[REDACTED]	-KL-[REDACTED]V-[REDACTED]
N. tabacum cv BY2	[REDACTED]	-KL-[REDACTED]V-[REDACTED]
N. tabacum cv SC58	[REDACTED]	-KL-[REDACTED]V-[REDACTED]
D. stramonium	[REDACTED]	[REDACTED]-RFE
L. esculentum	[REDACTED]	[REDACTED]-I
S. cerevisiae	SPFTSLYKVRDERTTDK[NEY]LPLKI	QFE---SKE
H. sapiens	T[PETFVO]SDERC	DMG-EV[F]SYLNI PGSEDVKLK[EE]
301		350
N. tabacum cv Xanthi	[REDACTED]	[AVF]K[Q]D[REP]
N. tabacum cv BY2	[REDACTED]	[AVF]R[Q]D[QPC]

Figure 13 (a)

20 / 21

N. tabacum cv SC58	[AVKQKQDDEP]	[A]
D. stramonium	[SAFRAEQHHEEQ]	[V]
L. esculentum	[APKQKETHEFP]	[R]
S. cerevisiae	[STAVLRLIEEEFPVGCGVD]	[YVAI]
H. sapiens	[ITGVINPDKYPSDSGVR]	[YYVAS]
351		
N. tabacum cv Xanthi	[-----]	[-----]
N. tabacum cv BY2	[-----]	[-----]
N. tabacum cv SC58	[-----]	[-----]
D. stramonium	[-----]	[N]
L. esculentum	[-----]	[T]
S. cerevisiae	[KISENEAMITYEVVNIAF]	[QEPHPRTYHNLEFHYD]
H. sapiens	[TGSDDEDESSQTFMYYV]	[VLFNPHFHAKH-KP]
401		
N. tabacum cv Xanthi	[VLSSVTVT]	[-----]
N. tabacum cv BY2	[VLNTVT]	[-----]
N. tabacum cv SC58	[VLNSVTVT]	[-----]
D. stramonium	[CMAPSNLN]	[-----]
L. esculentum	[CMENNLNLN]	[-----]
S. cerevisiae	[DFESTAVLDSINKSEYPYKVSIW]	[GPCIAKEYMKHDVIGD]
H. sapiens	[LLQKRPDPDEKY-YSSSIW]	[GDRIVERCDYMHGE]
451		
N. tabacum cv Xanthi	[MIVSAGLGA]	[GTSVLSNS]
N. tabacum cv BY2	[MIVTMIG]	[GTSVATGS]
N. tabacum cv SC58	[MIVDNAG]	[EDVNS]
D. stramonium	[IPEANGCNGG]	[TSVAAAS]
L. esculentum	[IPEMDGATV]	[GSMVATTS]
S. cerevisiae	[FYPALEISSATQ]	[EQTDIVYIDSELD]
H. sapiens	[MLEENMAMVFASTER]	[QRPTIYYMSGPAELMQQFQNPDFPPE]
501		
N. tabacum cv Xanthi	[-----]	[533]
N. tabacum cv BY2	[-----]	
N. tabacum cv SC58	[-----]	
D. stramonium	[-----]	
L. esculentum	[-----]	
S. cerevisiae	[-----]	
H. sapiens	[VEEQDASTLPVSCAWESGMKRHRAACASASINV]	

Figure 13 (b)

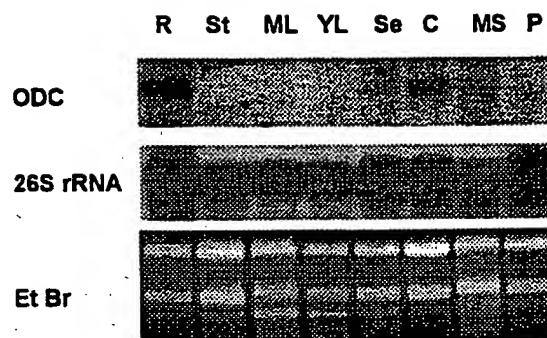
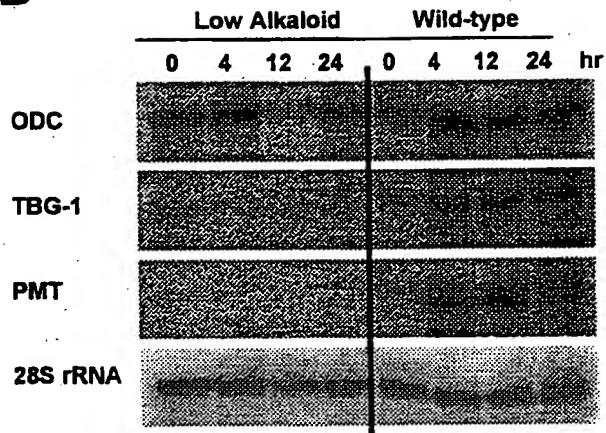
A**B**

Figure 14

SEQUENCE LISTING

<110> Timko P, Michael

<120> Regulation of Gene Expression in Tabacco for
Manipulation of Plant Growth and Secondary Metabolism

<130> 4981*239

<140>

<141>

<160> 26

<170> PatentIn Ver. 2.0

<210> 1

<211> 1120

<212> DNA

<213> Plant

<400> 1

ctgagttgac aagaacaatt cctggtaat cagatggatg aagataatacg aggtgggtgg 60
aatctataac caaaggcact gggtggatgt ctgtgcgagt tgcagaaaca attgaagggt 120
catttggaa atttggggcc atttcaaagg aaaaagaaaa gatgacttagt cattaataaa 180
tcaaattaaa ataaggctta gcgttaaaat caaaggaaat ggcaagcctg gtcctggag 240
caatgcttct gaggacagta gtaaaaaacaa tatcagacaa aaagtaaagt tgtattttt 300
agcttgagga taaagtatgt cattagttt gtgagagatt tggtgtccctc tacaatgatt 360
gttgaagtcc ctatttatag ctatacacag gaaacaaaat cctaggatca agccctctt 420
aaatgacaat aatggggta atgatgata tggatcgca tgacatgaat gccaaaattc 480
tccgcaacga ctatttattt aatattgagg aatattttt attaaataact atctggtgac 540
aagcattcgt ttgttccgt tgattacgtt gatttggga tctacttat accaaccgaa 600
gccgttgtcc ttgatctcg ctttcatatta attcatcttc cgtctgcctc cgatttcaca 660
agtcatgcac ccattcaatt atttaatgga aaccaatttt accctataca aatggtacat 720
cattcgtaa atacttact tggatataaa caatttgccc cgaggagtaa acagatgcga 780
agaaaagaaag cagacgattha aagaaatttt taaaaagga gagagaaatg aacacacaca 840
tgtactaata aaatttaggtt actactttac taataattgg acagagacta aattcatatt 900
ttagttccaa aatgtctcg ggatcacaac catgcacgtt gtaatgatt tttaactcta 960
ttatatcgag ttgcgcctc cactcctcg tggatataattt gtatataat gcatatgtgt 1020
ctattggag tggatcatcaa gcttcataa agtacaaatc gtaataacttgg ttgaaacata 1080
atactttctc ttctccaatt tggatgttt aattttgaaa 1120

<210> 2

<211> 3091

<212> DNA

<213> Plant

<400> 2

ctgaggatgac aagaacaatt cctggtaat cagatggatc aagataatag aggtgggtgg 60
aatctataac caaagcagct ggtttagtga ctgtgcagt tgcagaaaca attgaagggt 120
catttgtga atttggggcc atttcaaaagg aaaaagaaaa gatgacttag cattaataaa 180
tcaaattaaa ataaggctt gcgtttaaat caaagggaaat gcgaagcctg gctccctggag 240
caatgttct gaggacagta gtaaaaacaa tatcagacaa aaagtaagt tgtattattt 300
agcttgagga taaagtatgt cattagttt gtgagagatt tggtgtcctc tacaatgatt 360
gttgaagtcc ctatTTatAG ctatacacag gaaacaaaat cctaggatca agcccctt 420
aatgacaat aatggggta atgatgaata tgtagcggca tgacatgaat gccaaaattc 480
tccgcaacga ctatTTattt aatattgagg aatattttt attaaatact atctggtgc 540
aagcattcg ttcgttccgt tgattacgtt gatTTggga tctactctat accaaccgaa 600
gccgttgc ttgatctcg ctTTCattt attcatctc cgtctgcctc cgatttcaca 660
agtcatgcac ccattcaatt attaatgga aaccaattt accctataca aatggtacat 720
cattcgtaa atactttact tggatataaa caatTTGCC cgaggagtaa acagatgcg 780
agaaagaaag cagacgatta aagaaattt taaaaagga gagagaaatg aacacacaca 840
tgtactaata aaatttaggt actacttac taataattgg acagagacta aattcatatt 900
ttagttccaa aatgtctcg gcagTccaac catgcacggtt gtaatgatt tttaactcta 960
ttatATcgag ttgcgcctc cactcctcg tgcggaaattt gatatataat gcataatgtgt 1020
ctattggag tgcgttccaa gctttcataa agtacaaatc gtaataactt gtaaaacata 1080
atactttctc ttctccaatt tggatattt aattttgaaa atggaagtca tatctaccaa 1140
cacaatggc ttcattttttc tcaagagtgg tgccattccc atgaatggcc accataatgg 1200
cacttccaaa caccaaaacg gccacaagaa tggacttcc gaacaacaga acgggacaat 1260
cagccttgat aatggcaacg agctacttgg aaactccaat tgcgttgcgtt 1320
ttcagagttt agcgcattat ggccaggta gtaatggaa agaaactcaa atgcataattt 1380
aaagttaaaa ttgttaggtt aatataagga gttgatattc ttttagtgc taattaaaa 1440
ggaaaaagta tcaataaaat tcaaaaaatg gatagtaact tcgcattttt ctctacacat 1500
taatttgaaa taaatcgaaat ttgcgttgc aagcatttcc acttaaggtt gagaagttac 1560
tgcgttgcgg gaaatgttgc taccatgttgc tgcgttgc tgcgttgc tgcgttgc 1620
tacacatgttgc tccattttttt ttgcgttgc taccatgttgc tgcgttgc tgcgttgc 1680
gtacagtccatgttgc tccattttttt ttgcgttgc taccatgttgc tgcgttgc tgcgttgc 1740
aatggtgat ttccatCACAC tggatatttgc tgcgttgc taccatgttgc tgcgttgc tgcgttgc 1800
ccaaaaaagg ttgttaggttgc tccattttttt ttgcgttgc taccatgttgc tgcgttgc tgcgttgc 1860
tatcctacaa tggaaaaat ttgcgttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc 1920
caaacttctt ttactcacat aaaaaatgg ttgcgttgc taccatgttgc tgcgttgc tgcgttgc 1980
gaatactatt ttgttaggttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc tgcgttgc 2040
tatctcgctt ctaattttttt ttgcgttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc 2100
ttgcgttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc 2160
tgcgttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc 2220
ctgcacaacg agaataattt gatgttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc 2280
tattacttct taatccaaat tgcgttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc 2340
tttctaaaat aatataattt cagggtccacg aaaaatgg tttgcgttgc taccatgttgc tgcgttgc 2400
ggcagtagtgc aagccctaa ggccaggagg agttgtatgc acacaggctg aaagcattt 2460
gttgcgttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc 2520
tgtcaactat gctggacta ctgttccaaat tgcgttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc 2580
ttgcgttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc 2640
gaaaaaccaa cttcttttgc tttacttcc tggatatttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc 2700
atatgtcaat ttatTTGAT ttcagcgggttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc 2760
gaccagaaat tgacttcaag aatccagtaa atccaaatgttgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc 2820
agtccaaattt agcacccttc aagttctaca actctgtatgc taccatgttgc tgcgttgc taccatgttgc tgcgttgc 2880

cttttttccct attgtacttt atgttcttcg tcaaattta taattaactc ttttcaaattt 2940
 gtctttttt ttttcagatt cacaaggcag cattcatttt gccatcttc gccaagaagt 3000
 tgatcgagtc ttaatcaact gattaatgaa tactggtggt acaatcattg gaccaagatc 3060
 ataagtgaa agacgtattg tatgagaatt c 3091

<210> 3

<211> 353

<212> PRT

<213> Plant

<400> 3

Met	Glu	Val	Ile	Ser	Thr	Asn	Thr	Asn	Gly	Ser	Thr	Ile	Phe	Lys	Ser
1															15

Gly	Ala	Ile	Pro	Met	Asn	Gly	His	His	Asn	Gly	Thr	Ser	Lys	His	Gln
															30

Asn	Gly	His	Lys	Asn	Gly	Thr	Ser	Glu	Gln	Gln	Asn	Gly	Thr	Ile	Ser
35								40						45	

Leu	Asp	Asn	Gly	Asn	Glu	Leu	Leu	Gly	Asn	Ser	Asn	Cys	Ile	Lys	Pro
50								55						60	

Gly	Trp	Phe	Ser	Glu	Phe	Ser	Ala	Leu	Trp	Pro	Gly	Glu	Ala	Phe	Ser
65							70				75			80	

Leu	Lys	Val	Glu	Lys	Leu	Leu	Phe	Gln	Gly	Lys	Ser	Asp	Tyr	Gln	Asp
85								90						95	

Val	Met	Leu	Phe	Glu	Ser	Ala	Thr	Tyr	Gly	Lys	Val	Leu	Thr	Leu	Asp
100								105						110	

Gly	Ala	Ile	Gln	His	Thr	Glu	Asn	Gly	Gly	Phe	Pro	Tyr	Thr	Glu	Met
115								120						125	

Ile	Val	His	Leu	Pro	Leu	Gly	Ser	Ile	Pro	Asn	Pro	Lys	Val	Leu	
130								135						140	

Ile	Ile	Gly	Gly	Ile	Gly	Phe	Thr	Leu	Phe	Glu	Met	Leu	Arg	Tyr	
145								150						160	

Pro	Thr	Ile	Glu	Lys	Ile	Asp	Ile	Val	Glu	Ile	Asp	Asp	Val	Val	Val
165									170					175	

Asp	Val	Ser	Arg	Lys	Phe	Phe	Pro	Tyr	Leu	Ala	Ala	Asn	Phe	Asn	Asp
180									185					190	

Pro	Arg	Val	Thr	Leu	Val	Leu	Gly	Asp	Gly	Ala	Ala	Phe	Val	Lys	Ala

195

200

205

Ala Gln Ala Glu Tyr Tyr Asp Ala Ile Ile Val Asp Ser Ser Asp Pro
 210 215 220

Ile Gly Pro Ala Lys Asp Leu Phe Glu Arg Pro Phe Phe Glu Ala Val
 225 230 235 240

Ala Lys Ala Leu Arg Pro Gly Gly Val Val Cys Thr Gln Ala Glu Ser
 245 250 255

Ile Trp Leu His Met His Ile Ile Lys Gln Ile Ile Ala Asn Cys Arg
 260 265 270

Gln Val Phe Lys Gly Ser Val Asn Tyr Ala Trp Thr Thr Val Pro Thr
 275 280 285

Tyr Pro Thr Gly Val Ile Gly Tyr Met Leu Cys Ser Thr Glu Gly Pro
 290 295 300

Glu Ile Asp Phe Lys Asn Pro Val Asn Pro Ile Asp Lys Glu Thr Ala
 305 310 315 320

Gln Val Lys Ser Lys Leu Ala Pro Leu Lys Phe Tyr Asn Ser Asp Ile
 325 330 335

His Lys Ala Ala Phe Ile Leu Pro Ser Phe Ala Arg Ser Met Ile Glu
 340 345 350

Ser

<210> 4

<211> 711

<212> DNA

<213> Plant

<400> 4

gaattcaatg gagaaggaaa atatttccag tgtaaacaca agtgaatgaa gagaagccaa 60
 aataatctct atcattcaag ccttaggtgg agataaaaaa aattattttac tttcttatca 120
 aagtaatagg tgatcaacag ct当地cgtaaa acgtcattag gagaatatta taatctcttt 180
 tatgctgaag aaccacata aggaagatca taaaatacat gactttcaga tgacttcttg 240
 gagctttatt tt当地aaaggt ggcttagctgg tcagcaaaga ggtgctcgtc agatatcata 300
 aaattttact attatttgc ttaagaggaa gatggggcac acatgcttgc gacaaaagta 360
 agaggaagaa aggagacaga agaggaaaata gatttggggg gggggggggg ggtttcacaa 420
 tcaaagaaaa tt当地aaaat ggagagagaa atgagcacac acatataacta acaaatttt 480
 actaataattt gcaccgagac aaacttatattt ttttagttcca aatgtcagt ctaaccctgc 540

acgttgtaat gaattttaa ctattatatt atatcgagtt gcgcctcca ctcctcggtg 600
tccaaattgt atttaaatgc atagatgtt attggagtg tacagcaagc ttccggaaaa 660
tacaaaccat aataacttct cttttcaat ttgttagtt taattttgaa a 711

<210> 5

<211> 3129

<212> DNA

<213> Plant

<400> 5

gaattcaatg gagaaggaaa atattccag tgtaaacaca agtgaatgaa gagaagccaa 60
aataatctct atcattcaag ctttaggtgg agattaaaaa aattatttac ttttttatca 120
aagtaatagg tgatcaacag cttcgtaaa acgtcattag gagaatatta taatctttt 180
tatgctgaag aaccacata aggaatgca taaaatacat gactttcaga tgacttctg 240
gagctttatt tttaaagagt ggctagctgg tcagcaaaga ggtgctcgic agatatcata 300
aaattttact attatttgtt ttaagagggg gatggggcac acatgcttgt gacaaaagta 360
agaggaagaa aggagacaga agagggaaa gatttggggg gggggggggg ggtttcacaa 420
tcaaagaaaa ttttaaaat ggagagagaa atgagcacac acatatacta acaaaaattt 480
actaataatt gcaccgagac aaaccttat tttttttcca aaatgtcagt ctaaccctgc 540
acgttgtaat gaattttaa ctattatatt atatcgagtt gcgcctcca ctcctcggtg 600
tccaaattgt atttaaatgc atagatgtt attggagtg tacagcaagc ttccggaaaa 660
tacaaaccat aataacttct cttttcaat ttgttagtt taattttgaa aatggaagtc 720
atatctacca acacaaatgg ctctaccatc ttcaagaatg gtgcatttcc catgaacggc 780
caccaaaatg gcacttctga acacctcaac ggctaccaga atggcacttc caaacaccaa 840
aacggggcacc agaatggcac tttcgaacat cggaacggcc accagaatgg gacatccgaa 900
caacagaacg ggacaatcag ccatgacaat ggcaacgagc tactggaaag ctccgactct 960
attaagcctg gctggttttc agagtttgc gcattatggc caggttagta ctaagaaagc 1020
aactcaaatg catccgcctc ttgttgcata taaatataga gagctatcat acttttaggg 1080
actaactaaa aagggaaagat taticacagg acgaagttag cagtttaactt cgcatattat 1140
cagacgcatt aatttggaaat aatcgaattt tgcaggtgaa gcattctcac ttaaggttga 1200
gaagttacta ttccagggga agtctgatta ccaagatgtc atgcttttg aggttaattaa 1260
tattctaata cacatgctt aattttaaatg gatacttttta atttacattt agtttattgc 1320
atgtgcacgt acagtccacgca acttatgggaa aggttgcatttgc tttggatgaa gcaatttcaac 1380
atacagagaa tggggattt ccatacactg aaatgattgt tcatctacca cttggttcca 1440
tccccaaaccc aaaaaagggtt ttgatcatcg gcggaggaat tggtttaca ttattcgaaa 1500
tgcttcgtta tccttcaatc gaaaaatgg acatttgcatttgc gatcgatgac gtggtagttg 1560
atgtaaatc aacttctttt acccacataa agaaaatgtt ttagattgca attttttta 1620
tttttctaaa agaataaaata tattctctt tttttttta aaacaaaattt ctctttctta 1680
caggatccaa gaaaattttt cccttatctg gcagctaattt ttaacgatcc tcgtgttacc 1740
ctagttctcg gagatgggtgc gtatgtataa gtctcggtt atattttattt tcacttgatt 1800
tttacctttt tttgtggta attaatcata taccattgggt tctctttacc ttcaggagct 1860
gcattttgtaa aggctgcaca agcgggatata tttatgtgttata ttatagtgaa ctcttcgtat 1920
cccatttggta cgcttattactt atttaataacc aagactattt ttatataataa agtactaag 1980
aaactaattt aataatttaat aaacgttaact gtaattgtt tctaaaataa tatataataat 2040
ttcagggttcca gcaaaaagttt tttttggagag gccattttt gaggcagtag ccaaagccct 2100
taggccagga ggagttgtat gcacacagggc tgaaagcattt tggcttcata tgcattttat 2160
taagcaatc attgcttaact gtcgtcaagt cttaagggtt tctgtcaactt atgcttggac 2220
aaccgttcca acatatccca cgttattctt ttctctctt ctcttcgtt cttttcgtat 2280

gcaatgtaaa tttataaaaat tggaagtccg ttttactttt ctatagacgt agatcctaaa 2340
 attgtcaaga aatggagaat tgacttacaa gaaaaatcaa cttctttca tttactattc 2400
 ttttggta caaacattac ttattatttc gttctaaaat gaaaatttat ttttatattt 2460
 taaaataatt tagctttaaa ctttaattt tacttgtat attttaata aaaaagattt 2520
 atagtcaaat aaatgttgtg accatataaa aaccccgca ttttaagat cataagttc 2580
 agagtcaaac gagtaattt attttagt tgccggtgcg gagtcattt atgtcataaa 2640
 aattgaaacg gagtgagaac attttattt cgagtaact ttcaaggat tggtttaat 2700
 ttcaagtgt actgatcaat gatgtcttaa atatttgat ttcaaggat tggtttaat 2760
 tatgctctgc tctactgaag ggccagaagt tgacttcaag aatccagtaa atccaattga 2820
 caaagagaca actcaagtca agtccaaattt aggaccttc aagttctaca actctgatgt 2880
 aacttcatat ctcacaattt cttttccgt tttactgtat gttttcgatc aaattttata 2940
 actaacttctt ttcatattgt ctttttttc agattcacaa agcagcattc attttaccat 3000
 ctttcgcag aagtatgatc gagtcttaat caagtgaata atgaacactg gtgtacaaat 3060
 cattggacca agatcgagtc ttaatcaagt gaataaataa gtgaaatgag acgtattgt 3120
 ggagaattc 3129

<210> 6

<211> 375

<212> PRT

<213> Plant

<400> 6

Met	Glu	Val	Ile	Ser	Thr	Asn	Thr	Asn	Gly	Ser	Thr	Ile	Phe	Lys	Asn
1															
														15	

Gly	Ala	Ile	Pro	Met	Asn	Gly	His	Gln	Asn	Gly	Thr	Ser	Glu	His	Leu
														30	
20															

Asn	Gly	Tyr	Gln	Asn	Gly	Thr	Ser	Lys	His	Gln	Asn	Gly	His	Gln	Asn
														45	
35															

Gly	Thr	Phe	Glu	His	Arg	Asn	Gly	His	Gln	Asn	Gly	Thr	Ser	Glu	Gln
														60	
50															

Gln	Asn	Gly	Thr	Ile	Ser	His	Asp	Asn	Gly	Asn	Glu	Leu	Leu	Gly	Ser
														80	
65															

Ser	Asp	Ser	Ile	Lys	Pro	Gly	Trp	Phe	Ser	Glu	Phe	Ser	Ala	Leu	Trp
														95	
85															

Pro	Gly	Glu	Ala	Phe	Ser	Leu	Lys	Val	Glu	Lys	Leu	Leu	Phe	Gln	Gly
														110	
100															

Lys	Ser	Asp	Tyr	Gln	Asp	Val	Met	Leu	Phe	Glu	Ser	Ala	Thr	Tyr	Gly
														125	
115															

Lys	Val	Leu	Thr	Leu	Asp	Gly	Ala	Ile	Gln	His	Thr	Glu	Asn	Gly	Gly
														140	
130															

Phe Pro Tyr Thr Glu Met Ile Val His Leu Pro Leu Gly Ser Ile Pro
145 150 155 160

Asn Pro Lys Lys Val Leu Ile Ile Gly Gly Gly Ile Gly Phe Thr Leu
165 170 175

Phe Glu Met Leu Arg Tyr Pro Ser Ile Glu Lys Ile Asp Ile Val Glu
180 185 190

Ile Asp Asp Val Val Val Asp Val Ser Arg Lys Phe Phe Pro Tyr Leu
195 200 205

Ala Ala Asn Phe Asn Asp Pro Arg Val Thr Leu Val Leu Gly Asp Gly
210 215 220

Ala Ala Phe Val Lys Ala Ala Gln Ala Gly Tyr Tyr Asp Ala Ile Ile
225 230 235 240

Val Asp Ser Ser Asp Pro Ile Gly Pro Ala Lys Asp Leu Phe Glu Arg
245 250 255

Pro Phe Phe Glu Ala Val Ala Lys Ala Leu Arg Pro Gly Gly Val Val
260 265 270

Cys Thr Gln Ala Glu Ser Ile Trp Leu His Met His Ile Ile Lys Gln
275 280 285

Ile Ile Ala Asn Cys Arg Gln Val Phe Lys Gly Ser Val Asn Tyr Ala
290 295 300

Trp Thr Thr Val Pro Thr Tyr Pro Thr Gly Val Ile Gly Tyr Met Leu
305 310 315 320

Cys Ser Thr Glu Gly Pro Glu Val Asp Phe Lys Asn Pro Val Asn Pro
325 330 335

Ile Asp Lys Glu Thr Thr Gln Val Lys Ser Lys Leu Gly Pro Leu Lys
340 345 350

Phe Tyr Asn Ser Asp Ile His Lys Ala Ala Phe Ile Leu Pro Ser Phe
355 360 365

Ala Arg Ser Met Ile Glu Ser
370 375

<211> 1134

<212> DNA

<213> Plant

<400> 7

gctgtacaaa aggatgtctc aaatcatttgaatattaat tctgcaatca acaagaaata 60
ccccactatt aagaccatt atcaactggca caaaaattat gagatcatta aacatcttaa 120
acctgtccct atttggaaaga gtgtggatg ggagatgcct cccagggagt acctaaagct 180
gaatactgtat ggaagtttta acaaacaat tggaaagca gggattggag ggatttcag 240
agatgaagag ggaggctttg tcatacgctt ttcatggatgcct ataatctata ataacatcag 300
tgaaggcagaa ttgaaagcca tcaagtatgg gtgtgaatgg tgcaaataca aaggaatatc 360
aaacttcatt gtggaaactg actcgaggat gatctatgac atactacaga ccaaaaatct 420
aagcaacaac aagttgaaac aagagaccga gaaattaatg gagattctgg acacctgcag 480
gacacctgtt acccattgcc ttgcgcgaagc aaatcaagtgc gcaactgggt ttgctaaaga 540
ggccaccaga gctaacgaag gtatcactca tacagatttt agacaggtat caaaagcggc 600
caagggccct ttcttcattgg atatgtggca ggtcccttat tttagaattt gatatgaaaa 660
atctaattttttttttaag ttaattctgt gtatagtgag aggaaatcgt ctaatatgtat 720
tttttgcctt tagactcttc ctctccttag gtaaaaaggt agctccgagg taaggtttat 780
gttcccttca gtgtacccctt tttttgttta tataatagac atggatggg tccagctaaa 840
cccccaacac cacagggat agatacctgg gtgattgggtt ttttttttaa aaaaaaaaaac 900
tttactaata attgcacggc gacaaaactt atattttagt tccaaaatgac cagtccaaacc 960
atgcacgtt taatgatttt ttaactctat tatatcgagt tccgcccctcc actcctcggt 1020
gtccaaattt tatttaatgt catagatgt tttattgggat gtgtacatca agctttcaga 1080
aaatacaaac cataataactt tctttctcc aatttgctta gtttaattt gaaa 1134

<210> 8

<211> 3269

<212> DNA

<213> Plant

<400> 8

gctgtacaaa aggatgtctc aaatcatttgaatattaat tctgcaatca acaagaaata 60
ccccactatt aagaccatt atcaactggca caaaaattat gagatcatta aacatcttaa 120
acctgtccct atttggaaaga gtgtggatg ggagatgcct cccagggagt acctaaagct 180
gaatactgtat ggaagtttta acaaacaat tggaaagca gggattggag ggatttcag 240
agatgaagag ggaggctttg tcatacgctt ttcatggatgcct ataatctata ataacatcag 300
tgaaggcagaa ttgaaagcca tcaagtatgg gtgtgaatgg tgcaaataca aaggaatatc 360
aaacttcatt gtggaaactg actcgaggat gatctatgac atactacaga ccaaaaatct 420
aagcaacaac aagttgaaac aagagaccga gaaattaatg gagattctgg acacctgcag 480
gacacctgtt acccattgcc ttgcgcgaagc aaatcaagtgc gcaactgggt ttgctaaaga 540
ggccaccaga gctaacgaag gtatcactca tacagatttt agacaggtat caaaagcggc 600
caagggccct ttcttcattgg atatgtggca ggtcccttat tttagaattt gatatgaaaa 660
atctaattttttttttaag ttaattctgt gtatagtgag aggaaatcgt ctaatatgtat 720
tttttgcctt tagactcttc ctctccttag gtaaaaaggt agctccgagg taaggtttat 780
gttcccttca gtgtacccctt tttttgttta tataatagac atggatggg tccagctaaa 840
cccccaacac cacagggat agatacctgg gtgattgggtt ttttttttaa aaaaaaaaaac 900
tttactaata attgcacggc gacaaaactt atattttagt tccaaaatgac cagtccaaacc 960
atgcacgtt taatgatttt ttaactctat tatatcgagt tccgcccctcc actcctcggt 1020

gtccaaattt tattttaaatg catagatatg tttattggga gtttacatca agcttcaga 1080
 aaatacacaac cataataactt tctcttctcc aatttgctta gtttaattt gaaaatggaa 1140
 gtcatatcta ccaacacaaaa tggctctact atcttcaaga atgggtccat tcccatgaac 1200
 ggttaccaga atggacttc caaacaccaa aacggccacc agaatggcac ttccgaacat 1260
 cggaacgccc accagaatgg gatttccgaa cacaaaacg gccaccagaa tggcacttcc 1320
 gagcatcaga acggccatca gaatgggaca atcaggcatg acaacggcaa cgagctacag 1380
 ctactggaa gctccaactc tattaagctt gggttgggtt cagagtttag cgcattatgg 1440
 ccaggttagt actaagaaaag aaactcaaat gcacgtact ctgttattct gcttgcgt 1500
 taatttagat gatgggtttt gactaagcac tgagttaaa aataaaaagt ttaaagttaa 1560
 attgttacta tagagagcta tatctttagg aactaactaa aaaggaaaaaa ttatcacata 1620
 aaattggat gaagtaagca gttacttcg catattttc gacacattaa tttgaaataa 1680
 atcgaatttt gcaggtgaag cattctact taaggtttag aagttactat tccagggaa 1740
 gtctgattac caagatgtca tgctcttga ggtaattat taatactaat agtcaagctc 1800
 atgtatgatt atattttaaag tggtattttt cgtttatttt taattttattt cacgtgtac 1860
 tacagtgc aacatatggg aagggtctga ctttggatgg agcaattcaa cacacagaga 1920
 atggtgatt tccatacact gaaatgatg ttcatcttcc acttggttcc atccccaaacc 1980
 ctaaaaaggt tttgatcate ggcggaggaa ttgggtttac attattcgaa atgcttcgtt 2040
 atcctacaat cgaaaaaattt gacattgtt agatcgatga cgtggtagtt gatgttaagtc 2100
 aaacttcttt tactcacata aaaaaatgtt ttagattctt atttttctaa aagaattttaa 2160
 acaaaaattttt ccgttttaca ggtatctaga aaattttcc cttatcttgc tgctatttt 2220
 agcgatcctc gtgttaaccct agtccttga gatgggtcgtt atttgataat ctcgtttta 2280
 ttttatcttt tacttttattt ttatattttt ttacctttt tgggtgtgtt taatttccct 2340
 gccattgtt ctttttattt caggggtctgc atttgtaaag gcccacaag caggatatta 2400
 tgatgttattt atagtgactt ctctgtatcc cattggactt ctattactac ttaataccaa 2460
 gactattttt attaaataag ctactaataa acgttaactct gatagtttc taaaataata 2520
 taatttcagg tccagcaaaa gacttggttt agaggccattt ctttgaggca gtagccaaag 2580
 ccctaaggcc aggaggagtt gtatgcacac aggctgaaag catttggctt catatgcata 2640
 ttatattttt aatcattgtt aactgtcgatcc aagtctttaa gggctctgtc aactatgtt 2700
 ggactactgt tccaaacatattt ccaacgttattt ttctcttcc ttttccctata aatggaaag 2760
 ttttgattttt ataaattgtca agaaatggag aatcatttcc aaaaaaaaaacc aatattttt 2820
 cttttactct tcaagggtgtt ttatattttt ttatattgtt actgtatcaat tattttgattt 2880
 tcagcggtgtt gatgggttattt atgctctgtt ctactgtt accagaagttt gacttcaaga 2940
 atccagtaaa tccaaatttgc aagagacaa ctcaagtcaaa gtcaccaattt gtcacccatca 3000
 agttctacaa ctctgtatgtt aacttcatatcc tcaattttttt ttttcttattt gtactttatg 3060
 ttcttagtca aattttataa ttaacttctt tcaaaattgtt tttttttttc agattcacaa 3120
 agcagcatttcc atttttgcattt ctttgccttccag aagtatgttcc gaggctttaat caagtgtacta 3180
 atgaataactg gcggtacaat cattggacca agatcgatgtt ttaatcaagt gaataaataa 3240
 gtgaaatgctt acgttattgtt taagaattt 3269

<210> 9

<211> 381

<212> PRT

<213> Plant

<400> 9

Met Glu Val Ile Ser Thr Asn Thr Asn Gly Ser Thr Ile Phe Lys Asn

1

5

10

15

Gly Ala Ile Pro Met Asn Gly Tyr Gln Asn Gly Thr Ser Lys His Gln
20 25 30

Asn Gly His Gln Asn Gly Thr Ser Glu His Arg Asn Gly His Gln Asn
35 40 45

Gly Ile Ser Glu His Gln Asn Gly His Gln Asn Gly Thr Ser Glu His
50 55 60

Gln Asn Gly His Gln Asn Gly Thr Ile Ser His Asp Asn Gly Asn Glu
65 70 75 80

Ieu Gln Leu Leu Gly Ser Ser Asn Ser Ile Lys Pro Gly Trp Phe Ser
85 90 95

Glu Phe Ser Ala Leu Trp Pro Gly Glu Ala Phe Ser Leu Lys Val Glu
100 105 110

Lys Leu Leu Phe Gln Gly Lys Ser Asp Tyr Gln Asp Val Met Leu Phe
115 120 125

Glu Ser Ala Thr Tyr Gly Lys Val Leu Thr Leu Asp Gly Ala Ile Gln
130 135 140

His Thr Glu Asn Gly Gly Phe Pro Tyr Thr Glu Met Ile Val His Leu
145 150 155 160

Pro Leu Gly Ser Ile Pro Asn Pro Lys Lys Val Leu Ile Ile Gly Gly
165 170 175

Gly Ile Gly Phe Thr Leu Phe Glu Met Leu Arg Tyr Pro Thr Ile Glu
180 185 190

Lys Ile Asp Ile Val Glu Ile Asp Asp Val Val Val Asp Val Ser Arg
195 200 205

Lys Phe Phe Pro Tyr Leu Ala Ala Asn Phe Ser Asp Pro Arg Val Thr
210 215 220

Leu Val Leu Gly Asp Gly Ala Ala Phe Val Lys Ala Ala Gln Ala Gly
225 230 235 240

Tyr Tyr Asp Ala Ile Ile Val Asp Ser Ser Asp Pro Ile Gly Pro Ala
245 250 255

Lys Asp Leu Phe Glu Arg Pro Phe Phe Glu Ala Val Ala Lys Ala Leu
260 265 270

Arg Pro Gly Gly Val Val Cys Thr Gln Ala Glu Ser Ile Trp Leu His
275 280 285

Met His Ile Ile Lys Gln Ile Ile Ala Asn Cys Arg Gln Val Phe Lys
290 295 300

Gly Ser Val Asn Tyr Ala Trp Thr Thr Val Pro Thr Tyr Pro Thr Gly
305 310 315 320

Val Ile Gly Tyr Met Leu Cys Ser Thr Glu Gly Pro Glu Val Asp Phe
325 330 335

Lys Asn Pro Val Asn Pro Ile Asp Lys Glu Thr Thr Gln Val Lys Ser
340 345 350

Lys Leu Ala Pro Leu Lys Phe Tyr Asn Ser Asp Ile His Lys Ala Ala
355 360 365

Phe Ile Leu Pro Ser Phe Ala Arg Ser Met Ile Glu Ser
370 375 380

<210> 10

<211> 469

<212> DNA

<213> Plant

<400> 10

gtcgacacct gattccacaa gtcatgcacc cattcaatta tttaatggaa accaattttt 60
ccctgtacaa atggcacaaa tactttcctt ggataaaaaac aattttgcct aaggagtaaa 120
cagatgcgaa gtaagaaagc agacgactaa agaaaatttt aaaaaagggag agagaaatga 180
gcacacacac gtactaataa aatttagggta ctacttact aataattgga cagagactaa 240
attcatatatt tagtccaaa atgtctcgaaa cagtccaaacc atgcacgtt 300
ttaactctat tatctcgagt tgcgeccccc actccctgt gtccaaatgt tatataaatg 360
catatatgtc tattgggagt gtacagcgg 420
tgaaacataa tactttctct tctccaaattt gtttagttt attttgaaa 469

<210> 11

<211> 3001

<212> DNA

<213> Plant

<400> 11

gtcgacacct gattccacaa gtcatgcacc cattcaatta tttaatggaa accaattttt 60
ccctgtacaa atggcacaaa tactttcctt ggataaaaaac aattttgcct aaggagtaaa 120
cagatgcgaa gtaagaaagc agacgactaa agaaaatttt aaaaaagggag agagaaatga 180
gcacacacac gtactaataa aatttagggta ctacttact aataattgga cagagactaa 240
attcatatatt tagtccaaa atgtctcgaaa cagtccaaacc atgcacgtt 300

ttaactctat tatctcgagt tgcgccctcc actcctctgt gtccaagttg tatataaatg 360
 catatatgtc tattgggagt gtacagcgag ctttcataaa gtacaaaatca taataacttgt 420
 taaaacataa tacttctct tctccaattt gtttagttt atttgaaaa tggaaagtcat 480
 atctaccaac acaaatggct cgaccatctt caagaatggt gccattccca tgaatggcca 540
 ccagagtggc acttccaaac acctaaccgg ctaccagaac ggcaacttcca aacaccaaaa 600
 cggccaccat aatggcactt cccgaacatcg gaacggccac cagaatggg tttccgaaca 660
 cccaaacggc caccagaatg ggacttccga acatcgaaac ggcaccaga atgggatttc 720
 cgaacacccaa aacggccacc agaatgggac ttccgaacac caaaacggcc accagaatgg 780
 gacttccgaa caacagaacg ggacaatcg ccatgacaat ggcaacgagc tactggaaa 840
 ctccaaactt attaagctt gttggttttc agagtttagc gcattatggc caggtagta 900
 ctgagaaaga aactcaaatt catattnaa gttaaaattt ttaggctaataaagaatgt 960
 gatttcttt tagtgattaa taaaaaagg aaagagtatc aaataaattc caaaaaaaaatga 1020
 ccagtaactt cgcattttt tctacacatt aatttggaaat aaatcgaatt ttgcaggtga 1080
 agcattctcc cttaaagggtt aagttactt attcagggg aagtctgact accaagatgt 1140
 catgctcttt gaggtaaata atattctaat acacatgctt taatatgaat aaataacttt 1200
 aatttacttt tagtttattt cactgttacg tacagtccgc aacatatggg aagggtttga 1260
 ctttggatgg agcaattcaa cacacagaga atggggatt tccataactt gaaatgattt 1320
 ttcatcttcc acttgggttcc atcccaaacc caaaaaaaaatg ttgatcatc ggcggaggaa 1380
 ttgggtttac attattcgaa atgcttcgtt atccataat cggaaaaattt gacatttttg 1440
 aaatcgatga cgtggtagtt gatgttagtc aaatttcttt tactcacata aaaaaatgtat 1500
 ttagattgtt tcttttattt ttctaaaag aataaatata ttctcttta gttttaaaca 1560
 aaattctctt tcttacaggt atctagaaaaa tctttccctt atctcgccgc taattttat 1620
 gatecctcggt taaccctcgat tctcgagat ggtcgatatt tataatctcg tttttgtttt 1680
 atcttttattt ttatattcat ttaatttacc tttttgtgt tggttaattt acccgcttatt 1740
 ggttctctt catttcaggg gtcgcattt taaaggctgc acaagcagga tattatgatg 1800
 ctattatagt ggacttctt gatcccattt gtaacttattt actacttaat accaagacta 1860
 atcttatttga ataagctact aataaactgt aatttatttcaaaaataata taatttcagg 1920
 tccagaaaa gattttttt agaggccatt ctttgaggca gttagccaaag ccctaaggcc 1980
 aggaggagtt gtatgcacac aggccgaaag catttggctt catatgcata ttattaagca 2040
 aatcattgtt aactgtcgat aagtctttaa gggctctgc aactacgtt ggactactgt 2100
 tccaacatattt cccacgtatt ttctctctt ctctcttcat ctttggaaat tggaaatctt 2160
 gactactttc ttcccttga ttccctcggtt aaaggggcgt agatcataag atttcaaga 2220
 aataagataat gacgtccaaag aaaaactaaat ttctttcat ttactattt ttttggtgac 2280
 aaactttattt tattatttcg ttctaaagag aaaattttttttt ttttatattt aaaaataattt 2340
 tttttaaac ttttattttt attattata tctttaataa aaaaattata gtcaaaataaa 2400
 tattatggcc acactaaaca tccaagttt tggaaaccata agtttttagag ccaaatgagt 2460
 taattttttt ttgttatgcg ggtcgaggat caaattatgt cacaatgggat gtaatggagt 2520
 gagcaaaattt ttatattcgag taaactttca aggtattgtt tttttttt ttcactgtat 2580
 actaatcaat tatgtctcaa ccattttgtt ttcaactgtt ggtttttt tttttttt 2640
 tctactgaag ggccagaatg tgacttcaag aatccaataa atccaatttca caaagagaca 2700
 actcaagtca agtccaaattt agcacccttc aagttttaca attctgtat aacttcatat 2760
 ctaacaattt cttttctgt ttactgtat ttcaactgtt ctttggat tttttttt 2820
 tctcaaaattt tctttttttt tagattcaca aagcagcattt ctttggat tttttttt 2880
 gaagtatgtat cttttttttt tttttttt tttttttt tttttttt 2940
 aagatcgagt cttaatcaag tgaataaata agtgaaatgc cgacgttattt tatgagaatt 3000
 C
 3001

<211> 419

<212> PRT

<213> Plant

<400> 12

Met Glu Val Ile Ser Thr Asn Thr Asn Gly Ser Thr Ile Phe Lys Asn
1 5 10 15

Gly Ala Ile Pro Met Asn Gly His Gln Ser Gly Thr Ser Lys His Leu
20 25 30

Asn Gly Tyr Gln Asn Gly Thr Ser Lys His Gln Asn Gly His His Asn
35 40 45

Gly Thr Ser Glu His Arg Asn Gly His Gln Asn Gly Ile Ser Glu His
50 55 60

Gln Asn Gly His Gln Asn Gly Thr Ser Glu His Arg Asn Gly His Gln
65 70 75 80

Asn Gly Ile Ser Glu His Gln Asn Gly His Gln Asn Gly Thr Ser Glu
85 90 95

His Gln Asn Gly His Gln Asn Gly Thr Ser Glu Gln Gln Asn Gly Thr
100 105 110

Ile Ser His Asp Asn Gly Asn Glu Leu Leu Gly Asn Ser Asn Ser Ile
115 120 125

Lys Leu Gly Trp Phe Ser Glu Phe Ser Ala Leu Trp Pro Gly Glu Ala
130 135 140

Phe Ser Leu Lys Val Glu Lys Leu Leu Phe Gln Gly Lys Ser Asp Tyr
145 150 155 160

Gln Asp Val Met Leu Phe Glu Ser Ala Thr Tyr Gly Lys Val Leu Thr
165 170 175

Leu Asp Gly Ala Ile Gln His Thr Glu Asn Gly Gly Phe Pro Tyr Thr
180 185 190

Glu Met Ile Val His Leu Pro Leu Gly Ser Ile Pro Asn Pro Lys Lys
195 200 205

Val Leu Ile Ile Gly Gly Ile Gly Phe Thr Leu Phe Glu Met Leu
210 215 220

Arg Tyr Pro Thr Ile Glu Lys Ile Asp Ile Val Glu Ile Asp Asp Val

225	230	235	240
Val Val Asp Val Ser Arg Lys Ser Phe Pro Tyr Leu Ala Ala Asn Phe			
245	250	255	
Asn Asp Pro Arg Val Thr Leu Val Leu Gly Asp Gly Ala Ala Phe Val			
260	265	270	
Lys Ala Ala Gln Ala Gly Tyr Tyr Asp Ala Ile Ile Val Asp Ser Ser			
275	280	285	
Asp Pro Ile Gly Pro Ala Lys Asp Leu Phe Glu Arg Pro Phe Phe Glu			
290	295	300	
Ala Val Ala Lys Ala Leu Arg Pro Gly Gly Val Val Cys Thr Gln Ala			
305	310	315	320
Glu Ser Ile Trp Leu His Met His Ile Ile Lys Gln Ile Ile Ala Asn			
325	330	335	
Cys Arg Gln Val Phe Lys Gly Ser Val Asn Tyr Ala Trp Thr Thr Val			
340	345	350	
Pro Thr Tyr Pro Thr Gly Val Ile Gly Tyr Met Leu Cys Ser Thr Glu			
355	360	365	
Gly Pro Glu Val Asp Phe Lys Asn Pro Ile Asn Pro Ile Asp Lys Glu			
370	375	380	
Thr Thr Gln Val Lys Ser Lys Leu Ala Pro Leu Lys Phe Tyr Asn Ser			
385	390	395	400
Asp Ile His Lys Ala Ala Phe Ile Leu Pro Ser Phe Ala Arg Ser Met			
405	410	415	

Ile Glu Ser

<210> 13
<211> 1636
<212> DNA
<213> Plant

<400> 13
ggcacgagat cagatccaat tctttctgt gcttccttc tctgctctca aattcttcag 60
atctacaaag ttttcttcat ttcagaggg cagacatgga aactttcttg ttcacacctag 120
agtcagtcaa tgaaggccac cccgacaagc tctgcgacca ggtctcgat gcaattcttg 180

atgcttgctt agaacaggat ccagaaagca aggttgcatt tgaaacctgc acaaagacaa 240
 acatggttat ggtcttgga gagatcacaa ccaaggccac tggactat gagaagatag 300
 tgcgtgacac atgcagaggc attgggttca cctcagcaga tggccctt gacgctgaca 360
 actgcaaggt tcttgtaac atcgagcgc agagccctga cattgcccgg 420
 gtcatcttac caagaaacca gaagagattt gagctggta ccaaggtcac atggttgct 480
 atgccactga tgaaacccca gagctcatgc cccttacca tggcccttactaagctg 540
 gtgccaagct taccgaagt aggaagaaca agacttgcctt atggctcaga ccagatggca 600
 agacccaagt tactgttag tacaagaacg acaatggcgc catggccca attagagttc 660
 acactgttct catctcaact caacatgacg aaactgtcac aaacgaccag attgcccagg 720
 acttgaaaaga gcatgtgatc aaacctgtga tccatctca gtaccttgcata 780
 tcttccacccatca ggtcgcttcg tcacatggtagg accacacggg gatgctggac 840
 ttaccggcag gaaaattatc attgacacccatc acggaggctg ggggtccccat ggaggaggtg 900
 ctgttcagg aaaggaccct actaaggtagg acaggagttt tgcttatatt gtttagacagg 960
 cagcaaagag tgggtcgcc tcaggacttgc ctcgcccgtt tattgtgcag gtttctttag 1020
 ctatcggtgt ggctgaacca ctttccgtgt ttgttgcac ttacaagact ggaacaattc 1080
 cagacaagga tattttgact ctgatcaagg agaacttgc cttcaggccctt ggaatgtat 1140
 caatcaacct tgacttgttta agaggaggca acttcaggta ccagaagact gcagcttatg 1200
 gtcactttgg ccgtgatgac cccgacttct catggagac tgtcaaggcgc ctcaagccaa 1260
 aagcttaagt gaggtgttagc ctttggcca ttattttct tgcagaccaaa taaacaagct 1320
 tcacatcattt atgcatttgc ggcaggagaa gagaattttgt gtctccattt gaggattcta 1380
 tgagctctga gtcatttgcac attgttattt ttctttctt ttttttgcacc cttttcttgc 1440
 gtaccttattt tttatattttgt tactgttaag tagcagtgcattt ttaagtttc cctgttaagt 1500
 agcagtgttattt taagtttcc ctgttaagta gctggaaattha agtttccatg ttcttatcata 1560
 ttatatgtga acttgcattt tatcttgcattt ggtgaaagag tccttcaggaa aatagtttaa 1620
 aaaaaaaaaaaaaaaa 1636

<210> 14

<211> 390

<212> PRT

<213> Plant

<400> 14

Met Glu Thr Phe Leu Phe Thr Ser Glu Ser Val Asn Glu Gly His Pro

1 5 10

15

Asp Lys Leu Cys Asp Gln Val Ser Asp Ala Ile Leu Asp Ala Cys Leu
20 25 30Glu Gln Asp Pro Glu Ser Lys Val Ala Cys Glu Thr Cys Thr Lys Thr
35 40 45Asn Met Val Met Val Phe Gly Glu Ile Thr Thr Lys Ala Thr Val Asp
50 55 60Tyr Glu Lys Ile Val Arg Asp Thr Cys Arg Gly Ile Gly Phe Thr Ser
65 70 75 80

Ala Asp Val Gly Leu Asp Ala Asp Asn Cys Lys Val Leu Val Asn Ile

85

90

95

Glu Gln Gln Ser Pro Asp Ile Ala Gln Gly Val His Gly His Leu Thr
 100 105 110

Lys Lys Pro Glu Glu Ile Gly Ala Gly Asp Gln Gly His Met Phe Gly
 115 120 125

Tyr Ala Thr Asp Glu Thr Pro Glu Leu Met Pro Leu Thr His Val Trp
 130 135 140

Ala Thr Lys Leu Gly Ala Lys Leu Thr Glu Val Arg Lys Asn Lys Thr
 145 150 155 160

Cys Pro Trp Leu Arg Pro Asp Gly Lys Thr Gln Val Thr Val Glu Tyr
 165 170 175

Lys Asn Asp Asn Gly Ala Met Val Pro Ile Arg Val His Thr Val Leu
 180 185 190

Ile Ser Thr Gln His Asp Glu Thr Val Thr Asn Asp Gln Ile Ala Gln
 195 200 205

Asp Leu Lys Glu His Val Ile Lys Pro Val Ile Pro Ser Gln Tyr Leu
 210 215 220

Asp Glu Asn Thr Ile Phe His Leu Asn Pro Ser Gly Arg Phe Val Ile
 225 230 235 240

Gly Gly Pro His Gly Asp Ala Gly Leu Thr Gly Arg Lys Ile Ile Ile
 245 250 255

Asp Thr Tyr Gly Gly Trp Gly Ala His Gly Gly Ala Phe Ser Gly
 260 265 270

Lys Asp Pro Thr Lys Val Asp Arg Ser Gly Ala Tyr Ile Val Arg Gln
 275 280 285

Ala Ala Lys Ser Val Val Ala Ser Gly Leu Ala Arg Arg Cys Ile Val
 290 295 300

Gln Val Ser Tyr Ala Ile Gly Val Ala Glu Pro Leu Ser Val Phe Val
 305 310 315 320

Asp Thr Tyr Lys Thr Gly Thr Ile Pro Asp Lys Asp Ile Leu Thr Leu
 325 330 335

Ile Lys Glu Asn Phe Asp Phe Arg Pro Gly Met Met Ser Ile Asn Leu

340

345

350

Asp Leu Leu Arg Gly Gly Asn Phe Arg Tyr Gln Lys Thr Ala Ala Tyr
 355 360 365

Gly His Phe Gly Arg Asp Asp Pro Asp Phe Ser Trp Glu Thr Val Lys
 370 375 380

Val Leu Lys Pro Lys Ala
 385 390

<210> 15

<211> 1596

<212> DNA

<213> Plant

<400> 15

ggcacgaggg gaacaagaga aacatcatat tattgaatcc ctatgttttctt 60
 ttgatttcctt cctctcattt acctctctct tttcttcctt tgtttgatg gcccggccaa 120
 caatcatcggtt ccattttca gtccacaatt ggccggcgag 180
 cttctcttac agcggcgccg gccggcgaaa acggcaccag aaaagtcatc cctctctcaa 240
 gagatgcctt acaagatttc atgttatcaa tcataaccca aaaattacaa gatgagaaac 300
 aacctttta cgtgcttagac ttgggtgagg ttgtttctt tatggaccaa tggaaatctg 360
 ctctcccaa tatccgtcca ttttacgtctt taaaatgtaa ccctgaacccg tcgtttccctt 420
 caattttatc tgctatggc tcaaattttt attgtgttag ccgagctgaa attgagtatg 480
 ttttatctt tggcatttca cctgaccgtt ttgtttcgc aaatccatgc aaaccggaaat 540
 ccgatattat ttttgcagca aaagtgggg tgaatcttac aacctatgtat tctgaagacg 600
 aggtttacaa gatccgaaag catcacccga aatccgaaact cttgctccgc atcaagccca 660
 tgctcgacgg caacgcgaga tgcccaatgg gcccggaaata cggcgccgtt ccagaagaag 720
 tcgacccgct gctccggca gctcaagccg cccgtctcac cgtatccggc gtctcatcc 780
 acatcggtat cggagatgcc gattcaaagc cttatctcg cggccatagcc gcccgttaagg 840
 aagtgtttga aacagctgtt aactcgaaa tgcgtttata gactgttcta gacgtcgccg 900
 gcgggtttac atccggccac cagttcacaa ccggccgggtt cggccgttaaa tcagctttaa 960
 aacaacactt cgatgacgaa ccggaggttga caatcatagc tgaacccgggtt cgggtttttt 1020
 cagagacggc gtttactttt gcaacgcgaa ttataggaa aagagtgtt ggtgaatttga 1080
 gggagtttattt gattaacgc gggctgtacg gttcgatgaa ctgtgtactt tacgaccatg 1140
 cgacgggttac tgcaacgcgg ttagctgttc tgcgttatacg tagtaacgtt acctgcggcg 1200
 ggtcgaaaac gtttccgacg actgtgtttt ggcggacttg tgcgttatacg tagtaacgtt 1260
 taagggttatacg gacgtgttgcgg ttaatgtttt gttttttttt cctaataatgg 1320
 gtgcgttatacg taaagctgtt ggggttcaattt ttaatggatt taatacttcc gccattgtt 1380
 ctcacccgttcc tttttttttt ccaagctgtt gacccacccgtt tttttttttt tactaccgtt 1440
 gttttgtatgg tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1500
 aattttatatt ccaaatcagc ttgttatttcc tttttttttt tttttttttt tttttttttt 1560
 taattgtat tttctctaaa aaaaaaaaaaaaaaaa 1596

<210> 16

<211> 433

<212> PRT

<213> Plant

<400> 16

Met Ala Gly Gln Thr Ile Ile Val Ser Gly Leu Asn Pro Ala Ala Ile
1 5 10 15

Leu Gln Ser Thr Ile Gly Gly Ala Ser Pro Thr Ala Ala Ala Ala
20 25 30

Ala Glu Asn Gly Thr Arg Lys Val Ile Pro Leu Ser Arg Asp Ala Leu
35 40 45

Gln Asp Phe Met Leu Ser Ile Ile Thr Gln Lys Leu Gln Asp Glu Lys
50 55 60

Gln Pro Phe Tyr Val Leu Asp Leu Gly Glu Val Val Ser Leu Met Asp
65 70 75 80

Gln Trp Lys Ser Ala Leu Pro Asn Ile Arg Pro Phe Tyr Ala Val Lys
85 90 95

Cys Asn Pro Glu Pro Ser Phe Leu Ser Ile Leu Ser Ala Met Gly Ser
100 105 110

Asn Phe Asp Cys Ala Ser Arg Ala Glu Ile Glu Tyr Val Leu Ser Leu
115 120 125

Gly Ile Ser Pro Asp Arg Ile Val Phe Ala Asn Pro Cys Lys Pro Glu
130 135 140

Ser Asp Ile Ile Phe Ala Ala Lys Val Gly Val Asn Leu Thr Thr Tyr
145 150 155 160

Asp Ser Glu Asp Glu Val Tyr Lys Ile Arg Lys His His Pro Lys Ser
165 170 175

Glu Leu Leu Leu Arg Ile Lys Pro Met Leu Asp Gly Asn Ala Arg Cys
180 185 190

Pro Met Gly Pro Lys Tyr Gly Ala Leu Pro Glu Glu Val Asp Pro Leu
195 200 205

Leu Arg Ala Ala Gln Ala Ala Arg Leu Thr Val Ser Gly Val Ser Phe
210 215 220

His Ile Gly Ser Gly Asp Ala Asp Ser Asn Ala Tyr Leu Gly Ala Ile
225 230 235 240

Ala Ala Ala Lys Glu Val Phe Glu Thr Ala Ala Lys Leu Gly Met Ser
 245 250 255

 Lys Met Thr Val Leu Asp Val Gly Gly Phe Thr Ser Gly His Gln
 260 265 270

 Phe Thr Thr Ala Ala Val Ala Val Lys Ser Ala Leu Lys Gln His Phe
 275 280 285

 Asp Asp Glu Pro Glu Leu Thr Ile Ile Ala Glu Pro Gly Arg Phe Phe
 290 295 300

 Ala Glu Thr Ala Phe Thr Leu Ala Thr Thr Ile Ile Gly Lys Arg Val
 305 310 315 320

 Arg Gly Glu Leu Arg Glu Tyr Trp Ile Asn Asp Gly Leu Tyr Gly Ser
 325 330 335

 Met Asn Cys Val Leu Tyr Asp His Ala Thr Val Asn Ala Thr Pro Leu
 340 345 350

 Ala Val Leu Ser Asn Arg Ser Asn Val Thr Cys Gly Gly Ser Lys Thr
 355 360 365

 Phe Pro Thr Thr Val Phe Gly Pro Thr Cys Asp Ala Leu Asp Thr Val
 370 375 380

 Leu Arg Asp Tyr Gln Leu Pro Glu Leu Gln Val Asn Asp Trp Leu Val
 385 390 395 400

 Phe Pro Asn Met Gly Ala Tyr Thr Lys Ala Ala Gly Ser Asn Phe Asn
 405 410 415

 Gly Phe Asn Thr Ser Ala Ile Val Thr His Leu Ala Tyr Ser Tyr Pro
 420 425 430

 Ser

<210> 17
 <211> 2074
 <212> DNA
 <213> Plant

<400> 17
 tggttaactgg accgacgcga catttgcgt atatgtctta atcggcttag tcgctgacaa 60

catcatccac caagtcaaag ttccggaaatt catatcgaaa ctcatcatct tctatccgag 120
aaatgagggg actatctgtt tacggtaaaa accgagtcgt cccttcataat gactaatcg 180
gattagaaca taatggtcta aggttcatca ttataataac gagccatgtat atagagtttag 240
gttgtcaagc tcaagccccca gagagcgatc aatatcgaga tcgagccaag gttaaactcg 300
agaccttagag accgatcaat accggagaccg accaagtcaa ggtcgagctc gagaccaggaa 360
gaccggtaaa gattgagatc ggccaagatc gagatcgagc caagaaaattt aaaaactcg 420
atagccgtat ttagggagag aatctctgtt gaaatcacga cttgaatcag ggaaaaacta 480
attaattat ctatcatgtt atccccacta tgtatTTTtta attatactca aaatgggatt 540
ccccccactat attaagagtg gttatcattt gtaatggaga gacatacaca cattcattct 600
gacatataca gaaaatagag caaatactat ccttttttgg cttttgatat ttagtcatat 660
tgtttcttctt acccattgtt cttcactcaa tttggaggtt ataaaaacttgg aagggttaag 720
cttaacttagtc cattcgggtt gcattcattt cttttacaat aatttcgtca tcatttattt 780
attttcttcaa ttgtactaag ttataccacg tattttttaga actgcgtata aattcaactc 840
tatccatTTT tcgggttaaac accgaataca tagcacaataa gcaccctcaa ttgcaaaagt 900
ccaaagccaa gggttcattt ctttctgtt gaaatggagata gagaattggaa aatctaattt 960
agttatctaa atctttataa tttagccccc catataagaa aaaggaaaca aattaacttga 1020
agaacaatag cctcgcatag atttacccctt tccatataaaa ttttggat ttttactcaattt 1080
tttgcaaatg tgtctaaaat gataggactt gcaaaattttt atttacattt tcctactcct 1140
ctttaattttt caagaaaattt attttaagca ttctcgatTTt gctctgccc cttccgtccc 1200
ttgccccatctc tgactcggtt aggaccccttca ttgcaaaaaat ccaaacccaa ggaaccttcc 1260
atacattaca taagccacaa aatagtaact attaaaaactt accaataat cctcaaatac 1320
tcgcatTTT ttcataccta acacgtttac cttatcttctt cgtaatgcgtt ctacatttgc 1380
tagtgatata aaataccgaa ttaccacgc ggcaaccctt ccgtgtctt ccacggccc 1440
agagaatctc tttagcccccc aaatacgaaa attaacttctt agaattttt tttctgggta 1500
ttaccatgaa aataaaagaaa aagagaaaaag tcaagaaaattt taattgggctt aataactgggg 1560
tccactgccc agccacgcattt tttccctcta tataaaggctt cggtcacctt catgcaatc 1620
tcgctcactt cacagttgtt agtttacgtt tctttcttca attccatataa aagaaaaccctt 1680
tccgttaggt ttccgtccta ttttcttctt ttctacgtt cctttcttgc tatcaatatc 1740
tgtatgggtt ttttcttgcgtt cgaatttttag atttggggctt cctttatatac ctgtaaacctt 1800
ataatttctt gtttaaaacca aaaacttagc ttcttcttgc gtcaggggttgg ggatatttgg 1860
atcgtgttaag agtgtgttag aaggtgatca tcttttgcgtt cagttccctt tttgtttctt 1920
ttgagggggtt agccggggcc tccggcctcgg cgggtttaa tagcccccat ctattacaac 1980
cattgggcaaa aacatcattt aatctgtac aaagcaaaacc ctttaatttttgcgtt 2040
totattctttt qattttttaa cagaagaaga agag 2074

<210> 18

<211> 4321

<212> DNA

<213> Plant

<400> 18

tggtaactgg accgacgcga catttgcgt atatgtctta atcgggctag tcgctgacaa 60
catcatccac caagtcaaag ttccggaaatt catacggtt ctcatcatct tctatccgag 120
aaatgagggg actatctgta tacggtcaaa accgaggctg cccttcataat gactaatcga 180
gattagaaca taatggtcta aggttcatca ttataataac gagccatgat atagagttag 240
gttgtcaagc tcaagccccca gagagcgatc aatatcgaga tcgagccaag gttaaactcg 300
agaccttagag accgatcaat accgagacccg accaagtcaa ggtcgagctc gagaccccaga 360
gaccggtcaa gattgagatc ggccaagatc gagatcgagc caagaaaattt aaaaagtgcgtt 420

atagccgcat ttagggagag aatctctgcg gaaatcacga cttgaatcag ggaaaaacta 480
attaattaat ctatcatgtg atccccacta tgtatttttta attatactca aaatgggatt 540
cccccaactat attaagagt gttatcattt gtaatggaga gacatacaca cattcattct 600
gacatataca gaaaatagag caaatactat cctttttgg ctttgatata tttagtcata 660
tgtttcttct acccattgtt ctcaactcaa ttggaggtg ataaaacttg aaggtttaag 720
ctaactagtc catcggtt gcattcattt ctttacaat aattcgtca tcattttt 780
attttctcaa ttgtactaag ttataccacg tatttttaga actgcgtata aattcaactc 840
tatccatttt tcggtaaac accgaataca tagcacaata gcaccctcaa ttgcaaaagt 900
ccaaagccaa gggttcattt ctttctgaag aaatgagata gagaattgaa aatctaattt 960
agttatctaa atctttataa tttagccttc catataagaa aaaggaaaca aattaactga 1020
agaacaataag cctcgcatacg attacaccc tccatataaa tttgtttat actcaatttt 1080
tttgc当地at tgc当地aaat gataggactt gcaaattttt atttaacatt tcctactcct 1140
ctttaattttt caagaaatta attttaagca ttctcgattt gctctgccc cttcgcccc 1200
ttgccatctc tgactcggat agacacgc ttgcaaaaat ccaaacccaa ggaaccttcc 1260
atacattaca taagccacaa aatagtaact attaaaaact accaatatat cctcaaaatc 1320
tcgcgattat ttcataccta acacgtttac cttatcttct cgtaatgacg ctacattagt 1380
tagtgatata aaataccgaa ttaccacgc ggcaaccctc cgctgtctat ccacggccc 1440
agagaatctc ttagcccccc aaatacgaaa attaatttct agaattttt tttctggta 1500
ttaccatgaa aataaagaaa aagagaaaag tcaagaaatt taattgggt aatactgggg 1560
tccactgccc agccacgc tttccctcta tataaagcgt cgtcacctct catgcaaatac 1620
tcgctcaattt cacagttgtt agttcacgt tctttctca attccatataa aagaaaccct 1680
tccgttaggt tttccctcta tttctcttc ttctacgctt cctttctga tatcaatatac 1740
tgtatgggtt tttcttggtt cgaattttttag atttttttg ctttaatac ctgttaacctt 1800
ataattctct gtttaaacca aaaacttagc ttcttctgaa gtcagggtgg ggatattttg 1860
atcgtgtaaag agtggttag aaggtgatta tcttttgatt cagttccctt tttgcttctt 1920
ttgaggggggt agccggggcc tccgcctcg cgggtttaa tagcccccat ctattacaac 1980
cattgggcaa aaacatcatt aaatctgtac aaagcaacc cttaattttag ttttaattttc 2040
tgtattttt gattttttaa cagaagaaga agagatgccg gccctaggtt gttgcgtaga 2100
cgctactgtt tccccctc tccgcctatgc cttctctcg gatacgcttc ttcccgcgcc 2160
ggagttctt acctccggcg tacctcctac aaactccgcc gcccgttcca ttgggtctcc 2220
ggatctgtcc tctgtttgtt acgggggtcgta tgggtggga gctccttatt tctccgttaa 2280
ctctaacgga gatatctccg tccgaccaca tggtagggac acactcccc accagggaaat 2340
tgaccttctc aaggctgtga aaaaggccctc cgaccggaaa aattcagggg ggctcgggt 2400
tcagctgcctt ctgtgttttc gttccctga tgtctaaaa aaccgggtgg aatctctgca 2460
atcggctttt gatctcgctg ttcatccca gggctatggg gcccactacc aagggtttta 2520
tcccgtgaaa tgcaatcaag acaggttctg ggtggaaagat attgtcaaat tcgggtctc 2580
atcccggttc ggggttggaaag ctgggtctaa acccgagctc ctgttagcca tgagctgtct 2640
ctgcagggggc agtgcgtgagg gccttctcg ttcgaaatggt ttcaaggacg ctgagttacat 2700
ttcgcttctg tttgttgcaaa gaaagctcat gttaaacact gtaattttc ttgaacaaga 2760
ggaggagctt gacccgttga ttgtatataag ccgtaagatg gctgttcggc ccgttaattgg 2820
acttcgggtt aagctcgaga ccaagccattt aggccatttt ggatccactt ctggagaaaa 2880
aggtaagttt gggcttacaa cgaccggaaa tggctgtgtc gtgaagaagc tggaaagatc 2940
cggaatgtg gattgccttc agttgcgtca ttttcacatt ggatctcaga tcccttcaac 3000
ggcgttgctt gctgtgtgtt ttgggtggc tggctcaggatt tattgtgaat taatccgtct 3060
tgggtgcgggtt atgaagttca ttgatactgg aggtgggctc ggaatttgatt atgatgggtac 3120
taaatcatgtt gattcagatg tctctgttgg ctatggcattt caagaatacg cttccacagt 3180
tgtccaggcg gttcaatatag ttgcgaccg taagggctg aagcacccag tgatttgcag 3240
cgaaagtggc agggcaattt tttctcatca ctcaattctg attttcgaag ccgtgtctc 3300

ttcttagtcac tcatgttctt cttcacatct gtcttctggc ggctccaat ccatggcgga 3360
 gacgctcaat gaagatgccc ttgctgatta ccgcaattta tctgctgctg cagttcgtgg 3420
 agagtacgag acgtgtgtac ttactctga tcagttaaa cagagatgtg tggatcagtt 3480
 taaagaaggg tccttggta ttgaacatct tgctgctgtt gatacatct gtgattttgt 3540
 atcaaaggct atggggctg ctgatccat cccgacttac catgtgaatc tgtcaattt 3600
 cacttcaatt cctgattttt ggcctttgg tcaattgttt ccgattgttc caatacacccg 3660
 ttagatgaa aagcctgcag taagggaat attatccggac ttgacttgc acagtgtatgg 3720
 gaaggttcat aagttcattt gttggcgaatc aagcttgcag ctgcatttgcatgatgg 3780
 tggcgatggt ggtgggtatt atctggggat gttttgggt ggggcttatg aggaggcgct 3840
 cggaggactc cacaacctgt ttggtgacc aagcgtggtgc cgctgggtgc agagcgatag 3900
 cgctcacagc ttgcacatgt ctgcctccgt ccctggcccg tcctgcggcg acgtgcctcg 3960
 agcgatgcag cacgagcccg agctcatgtt cgagactctc aagcaccgtg cggaggaatt 4020
 cttggaacaa gaagaagaca aagggtggc cattgcattt ttggccagca gcttagctca 4080
 gtccttccat aacatgcattt accttgcgc gcctgcattt tgctgcttca ctgcagttac 4140
 tgctaacaac ggtggctata actactatta cagtgtatgaaatgcagcag atictgcata 4200
 aggggaggat gagatttgc accttgcac tgcttgaagt gttgtcgtag catctccagt 4260
 ttagtttgtt cgtcgaagtt gtctgtttt gaataatacc cttagttgtt gatgttttc 4320
 t 4321

<210> 19

<211> 720

<212> PRT

<213> Plant

<400> 19

Met Pro Ala Leu Gly Cys Cys Val Asp Ala Thr Val Ser Pro Pro Leu

1 5 10 15

Gly Tyr Ala Phe Ser Arg Asp Ser Ser Leu Pro Ala Pro Glu Phe Phe

20 25 30

Thr Ser Gly Val Pro Pro Thr Asn Ser Ala Ala Gly Ser Ile Gly Ser

35 40 45

Pro Asp Leu Ser Ser Ala Leu Tyr Gly Val Asp Gly Trp Gly Ala Pro

50 55 60

Tyr Phe Ser Val Asn Ser Asn Gly Asp Ile Ser Val Arg Pro His Gly

65 70 75 80

Thr Asp Thr Leu Pro His Gln Glu Ile Asp Leu Leu Lys Val Val Lys

85 90 95

Lys Ala Ser Asp Pro Lys Asn Ser Gly Gly Leu Gly Leu Gln Leu Pro

100 105 110

Leu Val Val Arg Phe Pro Asp Val Leu Lys Asn Arg Leu Glu Ser Leu

115 120 125

Gln Ser Ala Phe Asp Leu Ala Val His Ser Gln Gly Tyr Gly Ala His
130 135 140

Tyr Gln Gly Val Tyr Pro Val Lys Cys Asn Gln Asp Arg Phe Val Val
145 150 155 160

Glu Asp Ile Val Lys Phe Gly Ser Ser Phe Arg Phe Gly Leu Glu Ala
165 170 175

Gly Ser Lys Pro Glu Leu Leu Leu Ala Met Ser Cys Leu Cys Arg Gly
180 185 190

Ser Ala Glu Gly Leu Leu Val Cys Asn Gly Phe Lys Asp Ala Glu Tyr
195 200 205

Ile Ser Leu Ala Leu Val Ala Arg Lys Leu Met Leu Asn Thr Val Ile
210 215 220

Val Leu Glu Gln Glu Glu Leu Asp Leu Val Ile Asp Ile Ser Arg
225 230 235 240

Lys Met Ala Val Arg Pro Val Ile Gly Leu Arg Ala Lys Leu Arg Thr
245 250 255

Lys His Ser Gly His Phe Gly Ser Thr Ser Gly Glu Lys Gly Lys Phe
260 265 270

Gly Leu Thr Thr Thr Gln Ile Val Arg Val Val Lys Lys Leu Glu Glu
275 280 285

Ser Gly Met Leu Asp Cys Leu Gln Leu Leu His Phe His Ile Gly Ser
290 295 300

Gln Ile Pro Ser Thr Ala Leu Leu Ala Asp Gly Val Gly Glu Ala Ala
305 310 315 320

Gln Ile Tyr Cys Glu Leu Ile Arg Leu Gly Ala Gly Met Lys Phe Ile
325 330 335

Asp Thr Gly Gly Leu Gly Ile Asp Tyr Asp Gly Thr Lys Ser Cys
340 345 350

Asp Ser Asp Val Ser Val Gly Tyr Gly Ile Gln Glu Tyr Ala Ser Thr
355 360 365

Val Val Gln Ala Val Gln Tyr Val Cys Asp Arg Lys Gly Val Lys His
370 375 380

Pro Val Ile Cys Ser Glu Ser Gly Arg Ala Ile Val Ser His His Ser
385 390 395 400

Ile Leu Ile Phe Glu Ala Val Ser Ala Ser Ser His Ser Cys Ser Ser
405 410 415

Ser His Leu Ser Ser Gly Gly Leu Gln Ser Met Ala Glu Thr Leu Asn
420 425 430

Glu Asp Ala Leu Ala Asp Tyr Arg Asn Leu Ser Ala Ala Ala Val Arg
435 440 445

Gly Glu Tyr Glu Thr Cys Val Leu Tyr Ser Asp Gln Leu Lys Gln Arg
450 455 460

Cys Val Asp Gln Phe Lys Glu Gly Ser Leu Gly Ile Glu His Leu Ala
465 470 475 480

Ala Val Asp Ser Ile Cys Asp Phe Val Ser Lys Ala Met Gly Ala Ala
485 490 495

Asp Pro Ile Arg Thr Tyr His Val Asn Leu Ser Ile Phe Thr Ser Ile
500 505 510

Pro Asp Phe Trp Ala Phe Gly Gln Leu Phe Pro Ile Val Pro Ile His
515 520 525

Arg Leu Asp Glu Lys Pro Ala Val Arg Gly Ile Leu Ser Asp Leu Thr
530 535 540

Cys Asp Ser Asp Gly Lys Val Asp Lys Phe Ile Gly Gly Glu Ser Ser
545 550 555 560

Leu Gln Leu His Glu Leu Gly Ser Asn Gly Asp Gly Gly Tyr Tyr
565 570 575

Leu Gly Met Phe Leu Gly Gly Ala Tyr Glu Glu Ala Leu Gly Gly Leu
580 585 590

His Asn Leu Phe Gly Gly Pro Ser Val Val Arg Val Val Gln Ser Asp
595 600 605

Ser Ala His Ser Phe Ala Met Ser Arg Ser Val Pro Gly Pro Ser Cys
610 615 620

Ala Asp Val Leu Arg Ala Met Gln His Glu Pro Glu Leu Met Phe Glu
625 630 635 640

Thr Leu Lys His Arg Ala Glu Glu Phe Leu Glu Gln Glu Glu Asp Lys
645 650 655

Gly Leu Ala Ile Ala Ser Leu Ala Ser Ser Leu Ala Gln Ser Phe His
660 665 670

Asn Met Pro Tyr Leu Val Ala Pro Ala Ser Cys Cys Phe Thr Ala Val
675 680 685

Thr Ala Asn Asn Gly Gly Tyr Asn Tyr Tyr Ser Asp Glu Asn Ala
690 695 700

Ala Asp Ser Ala Thr Gly Glu Asp Glu Ile Trp Ser Tyr Cys Thr Ala
705 710 715 720

<210> 20
<211> 2118
<212> DNA
<213> Plant

<400> 20
gaattcctta tccggatttc tggtaacgcag actgtatatat ggagtcatct tctcctcgat 60
tcgggattaa aatttaggtga ctgggacac cctaaatctc ccaagtggcg actctgaaat 120
aaataaaacaa atcccgtttc gattgtcctt aaattggaaa aaactccctt gtaccctccc 180
gggtacggaa aaaggagggtg tacagcaatg accccaaact tttattgcta tacatttga 240
gaaatcaact ttagtcaaat ttatgggtga aattcaatgt ggtatgattt atattagtc 300
ggacttttagc agatgtggtc acttcaattt gccccaaaaa taatgtacag ggataataat 360
aaaaagtact agaaatttga gtcataaagg ttttcaatt ttacaaaaga tattaagata 420
cttattaaat caaatgtact ttattatgt aatagcatga aaaaacagcc tcattccgcct 480
gtcctcaccc cacaaaaagg agatagagaa aggaaactaa tcttattttaa ttttccacat 540
ataaaaattta tatccttgtt taaaatccccca aaaaaaaaaa atcaataacta attattttaa 600
ttaatcatc cgtataagaa agaagctaat taactgactt acaaaactgaa tagatagcac 660
aatagcactc tcaattacaa aaatccaaag ccgagggtca ttctttcat caagaaattt 720
gatagggaat ggaaaatata atttaattat ctgaatctt ataatttatac ttccatata 780
agaaaaaggaa aacaattaa ctgaagagca tatagctcg catagattt ctttctccat 840
atgggtgggg aaaccgacaa accgcaccaa ttgcataatt cgagtcaaac tgagggaaaa 900
aaaatttcgac tatgggttgg ttgtatgg tttattgttgg ggataaaaaaa tcgatcataa 960
ttgggttgg tttttttaa cttaagaaaag tcaaaccgaa accaaacaaa cccgacattt 1020
catatataaa ttttttagat atatttata tataatata cttgttgtga tgtaatttat 1080
aaatatttct taaaatatt cataattttt tcttttaaga tattatttcg tacttagaac 1140
ttttgaatgt ttcttactcc tcttttaattt gtttagaaatt aatttgaagg agtttgaattt 1200
tgctccgccc ccgctccgtc ccgttgccat ccctgactca ataggataac agcaatctcg 1260
attgcaaaaaa tccaaaccca aggaaccttc ccaacattac ataagctaca aagtagagta 1320

gtttattaaa taactaccaa tatataccta aattctcgcg attatttcat acctaacacg 1380
 ctacccat cttctcgtaa tgacgctaca ttagttggtg atataaaata ccgaatttgc 1440
 cacgcggcaa tcctccgctg tctatccacg gcccggagaga atctcttagc cccccaaaga 1500
 tgaaaattaa cttctagaat tttatTTTct ggttattacc atgaaaataaa taaaataaaa 1560
 aaaaagagaa aagtaaagat attaattgg gctaaaactg gggtccacgg cccagccacg 1620
 catttcctc ctatataaag cgctcgtaacc tctcatgca atctcgctca ctacacagg 1680
 gtagtttca cggtcttcc tcaattccca taacagaaac ccttccgtt ggttccgtc 1740
 ctatTTTcc tcatcttctc cgTTTccctc tctgaaatca atatctgtat ggtgttttc 1800
 ttgttgcata ttttagattt gttgtctt aatacctata accttaaatt ctctgtttaa 1860
 accaaaaact tagtttctc tgaagtcaagg gtggggattt ttggatcgtg taagagtgtg 1920
 ttagagggtg attatcttt gattcagttc ctTTTGTCTC tcttttgagg gggtagccgg 1980
 ggctcggcc tcggcgggtt ttaatagccc ccatcttata caactattgg gcaaaaacat 2040
 cattaaatct gtacaaaaca aacccttaat ttagtttaat ttctgttatt cattgattt 2100
 ttaacagaag aagaagag 2118

<210> 21

<211> 4368

<212> DNA

<213> Plant

<400> 21

gaattcctta tccggatttc tggtaacgcag actgtaatat ggagtcatct tctcctcgat 60
 tcgggattaa aattaggta cttgggacac cctaaatctc ccaagtggcg actctgaaat 120
 aaataaaacaa atcccgttc gattgtcctt aaattggaaa aaactccctt gtaccctccc 180
 gggtaacggaa aaaggaggtg tacagcaatg accaaaaact tttattgcta tacatTTGta 240
 ggaatcaact ttagtcaaat ttatgggtga aattcaatgt ggtatgatt atattaggtc 300
 ggactttacg agatgtggtc acttcaattt gcccggaaaaaa taatgtacag ggataataat 360
 aaaaagtact agaaatttga gtcataaaagc ttttcaatt ttacaaaaga tattaagata 420
 cttattaaat caaatgtact ttattaaatgt aatagcatga aaaaacagcc tcatccgcct 480
 gtcctcaccc cacaAAAAGG agatagagaa aggaaaactaa tcttatttaa ttttccacat 540
 ataaaattta tatccTTGta taaaatccccaa aaaaaaaaaa atcaataacta attattttaa 600
 ttaatcatc cgtataagaa agaagctaatt taactgactt acaaactgaa tagatagcac 660
 aatagcactc tcaattacaa aatccaaag ccgagggtca ttctttcat caagaaaatta 720
 gatagggaat ggaaaatata attaatttat ctgaatTTTataatttata cttccatata 780
 agaaaaagga aacaáattaa ctgaagagca tatagcctcg catagattta ccttctccat 840
 atgggtgggg aaaccgacaa accgcaccaa ttcgataatt cgagtcaaac tgagggaaaa 900
 aaaattcgcac tatggTTTgg tttgatttgg tttattttgg ggataaaaaaa tcgatcataa 960
 ttggTTTgg tttgattttaa ctaaaagaaag tcaaaacccgaa accaaaacaaa cccgacatta 1020
 catatataaaa ttttttagat atatttaata tataaaatata cttgttgcata tgtaatttat 1080
 aaaaatttct taaaatatttata cataattttta tctttaaga tattatttcg tacttagaac 1140
 ttttgaatgt ttcttactcc tctttaattt gtttagaaatt aatttgaagg agtttgaatt 1200
 tgctccgccc ccgctccgtc ccgttgcatt ccctgactca ataggataac agcaatctcg 1260
 attgcaaaaaa tccaaacccaa aggaacccctc ccaacattac ataagctaca aagttagagta 1320
 gtttattaaa taactaccaa tatataccta aattctcgcg attatttcat acctaacacg 1380
 ctacccat cttctcgtaa tgacgctaca ttagttggtg atataaaata ccgaatttgc 1440
 cacgcggcaa tcctccgctg tctatccacg gcccggagaga atctcttagc cccccaaaga 1500
 tgaaaattaa cttctagaat tttatTTTct ggttattacc atgaaaataaa taaaataaaa 1560
 aaaaagagaa aagtaaagat attaattgg gctaaaactg gggtccacgg cccagccacg 1620

catttccctc ctatataaaag cgtcgtcacc tctcatgcaa atctcgctca ctacacagtt 1680
 gttagttca cggtcttcc tcaattccca taacagaaaac cttccgtta ggttccgtc 1740
 ctattttcc tcacatcttc cgtttccctc tctgaaatca atatctgtat ggtgttttc 1800
 ttgttcaaat tttagatttg ttttgtctt aataacctata accttaaatt ctctgtttaa 1860
 accaaaaact tagcttcc tgaagtcaagg gtggggattt ttggatcgta taagagtgtg 1920
 tttagagggtg attatcttt gattcagttc ctttttgc tctttgagg gggtagccgg 1980
 ggcctcgccc tcggccgggt ttaatagccc ccatctatta caactattgg gcaaaaaacat 2040
 cattaaatct gtacaaaaca aacccttaat ttagtttaat tttctgttatt cattgatttt 2100
 ttaacagaag aagaagagat gccggcccta gggtgtgtg tagatgctgc tggtgttcc 2160
 cctcctctca gctatgcctt ctctcggtt agctcttcc ccgcgccggc gttctttgcc 2220
 tccggcgtac ctccgacaaa ttctgcccgt gcttccattt ggtctccggc tttgtcgct 2280
 gctttatacg gggtcgtatgg gtggggagct ctttatttct ctgttaactc taatggagat 2340
 atctccgtcc gaccacacgg tacggacact ctccctcacc agggaaattga ctttctcaag 2400
 gtcgtaaaaa aggcccccga cccgaaaaat tcaggtggc ttgggcttca gtcgcctt 2460
 gttgttcgtc tccctgtatgt gttaaaaac cggttgaat ctctgcaatc ggcttttgc 2520
 ctgcgggttc attcccgagg ctatggggcc cactaccaag gtgttatcc cgtgaaatgc 2580
 aatcaagaca gggtcggtt ggaagatatac gtgaaattcg ggtcgccatt ccgggtcggg 2640
 ttggaagccg ggtctaaacc cgagctctg tttagccatga gctgtctctg caagggcagt 2700
 gctgagggcc ttctcgttt caatggttt aaggacgctg agtacatttgc gcttgctt 2760
 gttgcaagaa agctcatgtt aaacactgtt attgtgtttt aacaagagga ggagcttgac 2820
 ctgtgttattt atataagcca taagatggct gttcggccctg taattggact tcgggctaag 2880
 ctcaggacca agcattcagg ccatttttga tccacttctg gagaaaaaagg taagtttggg 2940
 cttacaacga cccaaattgt tcgtgtgggt aagaagctag aagaatccgg aatgctggat 3000
 tgtcttcagt tgctgcattt tcaatttggta tctcagatcc cttctacggg gttgcttagt 3060
 gatggagttt gtgaggccgc tcaagattt tctgttttttgc agcgggtatg 3120
 aagttcattt atattggagg tgggcttggta attgattatg atggacttatac atcatgcgt 3180
 tctgtgttctt ctgttggcta tggcattcaaa gaatatgcct ccgcagttgt tcaggcgtt 3240
 caatatgtat gcgaccgtaa ggggtgtaaa cacccagtga tctgcagcga aagtggcagg 3300
 gcaattgtt ctcatcaactt aattctgatt ttcgaagccg tgcgtgttcc tagtcactca 3360
 tgggttttcc cacatctgtc ttctgggtggc ctccaaatcca tggcggagac gctcaacgaa 3420
 gatgcccctt ctgattaccg caattttatct gctgctgcag ttctgggaga gtatgagaca 3480
 tctgtacttt actctgtatca gttaaaaacag agatgtgtgg atcagtttaa agaagggtcc 3540
 ttgggttattt aacatcttgc tgctgttgcgat agcatctgttgc attttgtatc aaaggctatg 3600
 ggggtctgttgc atcctgtccg cacttaccat gtgaatctgtt caattttccat ttcaatttcc 3660
 gatttttggg ctttgggtca attgtttccg attgttccaa ttccaccgtt agatgaaaag 3720
 cctgcagtgaa ggggaatattt atcggacttta acttgtgaca gtgatggaa ggttggataag 3780
 ttcattgggtt gcgaatcaag ctgcggctt catgaattgg gaagtaatgg cgatgggtt 3840
 ggttatttttcc tgggatgtt ttgggtggg gcttattggagg aggccgtccgg aggactccac 3900
 aacctgttttgc tgggaccaag tgcgtgtccg tgggtgcaga gcatagcgc tcacagctt 3960
 gccatgactc gctccgtcccc tggccgtct tgcgtgtatg tgctccgagc gatgcagcac 4020
 gagcccgagc tcatgttgcgacttcaag caccgtgcgg aggaattttt ggaacaagaa 4080
 gatgacaaag ggctggctgt tgaatcttttgc cccagcagcg tagctcagtc cttccataac 4140
 atgccttacc ttgtggccccc ttcatcttgc cgcttcaacttgc ctgtacttgc taacaatgg 4200
 ggctataattt actattacag tggatgagaat gcaagcagatt ctgttacagg ggaggatgag 4260
 atttggtctt attgttgcacttgc ttggatgttttgc tgcgttgc tccacttgc agtttgcgt 4320
 cgagggttgc tgggttgc tgggttgc taataccctt agttgggtat gtttttgc 4368

<211> 721

<212> PRT

<213> Plant

<400> 22

Met Pro Ala Leu Gly Cys Cys Val Asp Ala Ala Val Val Ser Pro Pro
1 5 10 15

Leu Ser Tyr Ala Phe Ser Arg Asp Ser Ser Leu Pro Ala Pro Glu Phe
20 25 30

Phe Ala Ser Gly Val Pro Pro Thr Asn Ser Ala Ala Ala Ser Ile Gly
35 40 45

Ser Pro Asp Leu Ser Ser Ala Leu Tyr Gly Val Asp Gly Trp Gly Ala
50 55 60

Pro Tyr Phe Ser Val Asn Ser Asn Gly Asp Ile Ser Val Arg Pro His
65 70 75 80

Gly Thr Asp Thr Leu Pro His Gln Glu Ile Asp Leu Leu Lys Val Val
85 90 95

Lys Lys Ala Ser Asp Pro Lys Asn Ser Gly Gly Leu Gly Leu Gln Leu
100 105 110

Pro Leu Val Val Arg Phe Pro Asp Val Leu Lys Asn Arg Leu Glu Ser
115 120 125

Leu Gln Ser Ala Phe Asp Leu Ala Val His Ser Gln Gly Tyr Gly Ala
130 135 140

His Tyr Gln Gly Val Tyr Pro Val Lys Cys Asn Gln Asp Arg Phe Val
145 150 155 160

Val Glu Asp Ile Val Lys Phe Gly Ser Pro Phe Arg Phe Gly Leu Glu
165 170 175

Ala Gly Ser Lys Pro Glu Leu Leu Ala Met Ser Cys Leu Cys Lys
180 185 190

Gly Ser Ala Glu Gly Leu Leu Val Cys Asn Gly Phe Lys Asp Ala Glu
195 200 205

Tyr Ile Ser Leu Ala Leu Val Ala Arg Lys Leu Met Leu Asn Thr Val
210 215 220

Ile Val Leu Glu Gln Glu Glu Leu Asp Leu Val Ile Asp Ile Ser

225	230	235	240
His Lys Met Ala Val Arg Pro Val Ile Gly Leu Arg Ala Lys Leu Arg			
245	250	255	
Thr Lys His Ser Gly His Phe Gly Ser Thr Ser Gly Glu Lys Gly Lys			
260	265	270	
Phe Gly Leu Thr Thr Gln Ile Val Arg Val Val Lys Lys Leu Glu			
275	280	285	
Glu Ser Gly Met Leu Asp Cys Leu Gln Leu Leu His Phe His Ile Gly			
290	295	300	
Ser Gln Ile Pro Ser Thr Gly Leu Leu Ala Asp Gly Val Gly Glu Ala			
305	310	315	320
Ala Gln Ile Tyr Cys Glu Leu Val Arg Leu Gly Ala Gly Met Lys Phe			
325	330	335	
Ile Asp Ile Gly Gly Leu Gly Ile Asp Tyr Asp Gly Thr Lys Ser			
340	345	350	
Cys Asp Ser Asp Val Ser Val Gly Tyr Gly Ile Gln Glu Tyr Ala Ser			
355	360	365	
Ala Val Val Gln Ala Val Gln Tyr Val Cys Asp Arg Lys Gly Val Lys			
370	375	380	
His Pro Val Ile Cys Ser Glu Ser Gly Arg Ala Ile Val Ser His His			
385	390	395	400
Ser Ile Leu Ile Phe Glu Ala Val Ser Ala Ser Ser His Ser Cys Ser			
405	410	415	
Ser Ser His Leu Ser Ser Gly Gly Leu Gln Ser Met Ala Glu Thr Leu			
420	425	430	
Asn Glu Asp Ala Leu Ala Asp Tyr Arg Asn Leu Ser Ala Ala Ala Val			
435	440	445	
Arg Gly Glu Tyr Glu Thr Cys Val Leu Tyr Ser Asp Gln Leu Lys Gln			
450	455	460	
Arg Cys Val Asp Gln Phe Lys Glu Gly Ser Leu Gly Ile Glu His Leu			
465	470	475	480
Ala Ala Val Asp Ser Ile Cys Asp Phe Val Ser Lys Ala Met Gly Ala			

485

490

495

Ala Asp Pro Val Arg Thr Tyr His Val Asn Leu Ser Ile Phe Thr Ser
500 505 510

Ile Pro Asp Phe Trp Ala Phe Gly Gln Leu Phe Pro Ile Val Pro Ile
515 520 525

His Arg Leu Asp Glu Lys Pro Ala Val Arg Gly Ile Leu Ser Asp Leu
530 535 540

Thr Cys Asp Ser Asp Gly Lys Val Asp Lys Phe Ile Gly Gly Glu Ser
545 550 555 560

Ser Leu Pro Leu His Glu Leu Gly Ser Asn Gly Asp Gly Gly Tyr
565 570 575

Tyr Leu Gly Met Phe Leu Gly Gly Ala Tyr Glu Glu Ala Leu Gly Gly
580 585 590

Leu His Asn Leu Phe Gly Gly Pro Ser Val Val Arg Val Val Gln Ser
595 600 605

Asp Ser Ala His Ser Phe Ala Met Thr Arg Ser Val Pro Gly Pro Ser
610 615 620

Cys Ala Asp Val Leu Arg Ala Met Gln His Glu Pro Glu Leu Met Phe
625 630 635 640

Glu Thr Leu Lys His Arg Ala Glu Glu Phe Leu Glu Gln Glu Asp Asp
645 650 655

Lys Gly Leu Ala Val Glu Ser Leu Ala Ser Ser Val Ala Gln Ser Phe
660 665 670

His Asn Met Pro Tyr Leu Val Ala Pro Ser Ser Cys Arg Phe Thr Ala
675 680 685

Ala Thr Asp Asn Asn Gly Gly Tyr Asn Tyr Tyr Ser Asp Glu Asn
690 695 700

Ala Ala Asp Ser Ala Thr Gly Glu Asp Glu Ile Trp Ser Tyr Cys Thr
705 710 715 720

Ala

<210> 23

<211> 2695

<212> DNA

<213> Plant

<400> 23

ttcacgttctt cttctcaatt cccataaaaag aaacccttcc gtaggttcc cgcccttattt 60
tctcttc tacgcttcctt ctctgatataaatatctgt atgggtttt tcttgcga 120
attttagatt tggttgccctttaataacctg taaccttata attctctgtt taaaccaaaa 180
acttagcttc ttctgaagtc aggggtggga tatttggatc gtgtaaagatgtgttagaaag 240
gtgattatct tttgatttcag ttccctttt gcttctttt aggggttagc cggggcctcg 300
gcctcggcgg gtttaataag cccccatcta ttacaaccat tggcaaaaa catcattaaa 360
tctgtacaaa gcaaaccctt aatttagttt aattttctgtt attcttgcgtatctttaacag 420
aagaagaaga gatgccggcc cttagttgtt gcgttagacgc tactgtttcc cctccctctcg 480
gctatgcctt ctctcggat agctcttc ccgcgcggg gtttttacc tccggcgtac 540
ctcctacaaa ctccgcgcgc ggttccattt ggttccggg tctgtctctt gctttgtacg 600
gggtcgatgg gtggggagat ccttatttctt ccgttaactc taacggagat atctccgtcc 660
gaccacatgg tacggacaca ctccccccacc agggaaatttga ccttctcaag gtcgtgaaaa 720
aggcctccga cccgaaaaat tcaggggggc tcgggcttca getgccttctt gttgttcgt 780
tccctgtatgt gctaaaaaac cggttggat ctctgeaattc ggcttttgcgttcc 840
attcccagggtt ctagggggcc cactaccaag gtgttatcc cgtgaaatgc aatcaagaca 900
ggtctgttgtt ggaagatattt gtcatttgcgttccggg ttggaaagctg 960
ggtctaaacc cgagctccgtt ttagccatga gctgtctctg cagggcagt gctgaggggcc 1020
ttctcggtt caatggtttca aaggacgtt agtacatttgcgttcc 1080
agctcatgtt aaacactgtt attgttcttgcgtt aacaagagga ggagcttgcatttgcgttcc 1140
atataagccg taagatggat gtcggcccg taattggact tcgggcttaag ctcaggacca 1200
agcattcagg ccatttttggatccacttctg gagaaaaaagg taagtttggg cttacaacga 1260
cccaaattgt tcgttagtgcgtt aagaagctgg aagaatccgg aatgtctggat tgccttcgtt 1320
tgctgcattt tcacatttggatccacttctg cttcaacggc gttgttgcgttcc 1380
gtgaggctgc tcagatttgcgttcaatggat tccgttcc 1440
atactggagg tgggctcgatccacttctg tggacttgcgttcc 1500
ctgttggcttcaatggatccacttctg tggcatttgcgttcc 1560
gcgaccgtaa gggcggttccacttctg tttgcgttcc 1620
ctcatcacttccatttgcgttccacttctg tggacttgcgttcc 1680
cacatctgttccacttctg tggcatttgcgttcc 1740
ctgattaccgcgttccacttctg tggacttgcgttcc 1800
actctgtatgttccacttctg tggacttgcgttcc 1860
aacatcttgcgttccacttctg tggacttgcgttcc 1920
atccatccgttccacttctg tggacttgcgttcc 1980
cctttggcttccacttctg tggacttgcgttcc 2040
ggggatatttgcgttccacttctg tggacttgcgttcc 2100
gcgaaatccgttccacttctg tggacttgcgttcc 2160
tgggatgttccacttctg tggacttgcgttcc 2220
gtggaccaatccgttccacttctg tggacttgcgttcc 2280
gctccgtcccttccacttctg tggacttgcgttcc 2340
tcatgttccacttctg tggacttgcgttcc 2400
ggctggccat tgcgttccacttctg tggacttgcgttcc 2460
ttgtggcgcccttccacttctg tggacttgcgttcc 2520

actattacag tgatgagaat gcagcagatt ctgctacagg ggaggatgag atttggtcct 2580
 attgcactgc ttgaagtgtt gtcgttagcat ctccagttt agttgtcgt cgaagttgtc 2640
 tgttttgaa taataccctt agttggtgat gttttctaa aaaaaaaaaaaa aaaaaa 2695

<210> 24

<211> 720

<212> PRT

<213> Plant

<400> 24

Met	Pro	Ala	Leu	Gly	Cys	Cys	Val	Asp	Ala	Thr	Val	Ser	Pro	Pro	Leu
1															15

Gly	Tyr	Ala	Phe	Ser	Arg	Asp	Ser	Ser	Leu	Pro	Ala	Pro	Glu	Phe	Phe
															30
									20		25				

Thr	Ser	Gly	Val	Pro	Pro	Thr	Asn	Ser	Ala	Ala	Gly	Ser	Ile	Gly	Ser
															45
									35		40				

Pro	Asp	Leu	Ser	Ser	Ala	Leu	Tyr	Gly	Val	Asp	Gly	Trp	Gly	Ala	Pro
															50
									55			60			

Tyr	Phe	Ser	Val	Asn	Ser	Asn	Gly	Asp	Ile	Ser	Val	Arg	Pro	His	Gly
															80
									65		70		75		

Thr	Asp	Thr	Leu	Pro	His	Gln	Glu	Ile	Asp	Leu	Leu	Lys	Val	Val	Lys
															95
									85		90				

Lys	Ala	Ser	Asp	Pro	Lys	Asn	Ser	Gly	Gly	Leu	Gly	Leu	Gln	Leu	Pro
															100
									105			110			

Leu	Val	Val	Arg	Phe	Pro	Asp	Val	Leu	Lys	Asn	Arg	Leu	Glu	Ser	Leu
															115
									120			125			

Gln	Ser	Ala	Phe	Asp	Leu	Ala	Val	His	Ser	Gln	Gly	Tyr	Gly	Ala	His
															130
									135			140			

Tyr	Gln	Gly	Val	Tyr	Pro	Val	Lys	Cys	Asn	Gln	Asp	Arg	Phe	Val	Val
															145
									150		155		160		

Glu	Asp	Ile	Val	Lys	Phe	Gly	Ser	Ser	Phe	Arg	Phe	Gly	Leu	Glu	Ala
															165
									165		170		175		

Gly	Ser	Lys	Pro	Glu	Leu	Leu	Ala	Met	Ser	Cys	Leu	Cys	Arg	Gly	
															180
									185			190			

Ser	Ala	Glu	Gly	Leu	Leu	Val	Cys	Asn	Gly	Phe	Lys	Asp	Ala	Glu	Tyr
															195
									200			205			

Ile Ser Leu Ala Leu Val Ala Arg Lys Leu Met Leu Asn Thr Val Ile
210 215 220

Val Leu Glu Gln Glu Glu Glu Leu Asp Leu Val Ile Asp Ile Ser Arg
225 230 235 240

Lys Met Ala Val Arg Pro Val Ile Gly Leu Arg Ala Lys Leu Arg Thr
245 250 255

Lys His Ser Gly His Phe Gly Ser Thr Ser Gly Glu Lys Gly Lys Phe
260 265 270

Gly Leu Thr Thr Thr Gln Ile Val Arg Val Val Lys Lys Leu Glu Glu
275 280 285

Ser Gly Met Leu Asp Cys Leu Gln Leu Leu His Phe His Ile Gly Ser
290 295 300

Gln Ile Pro Ser Thr Ala Leu Leu Ala Asp Gly Val Gly Glu Ala Ala
305 310 315 320

Gln Ile Tyr Cys Glu Leu Ile Arg Leu Gly Ala Gly Met Lys Phe Ile
325 330 335

Asp Thr Gly Gly Leu Gly Ile Asp Tyr Asp Gly Thr Lys Ser Cys
340 345 350

Asp Ser Asp Val Ser Val Gly Tyr Gly Ile Gln Glu Tyr Ala Ser Thr
355 360 365

Val Val Gln Ala Val Gln Tyr Val Cys Asp Arg Lys Gly Val Lys His
370 375 380

Pro Val Ile Cys Ser Glu Ser Gly Arg Ala Ile Val Ser His His Ser
385 390 395 400

Ile Leu Ile Phe Glu Ala Val Ser Ala Ser Ser His Ser Cys Ser Ser
405 410 415

Ser His Leu Ser Ser Gly Gly Leu Gln Ser Met Ala Glu Thr Leu Asn
420 425 430

Glu Asp Ala Leu Ala Asp Tyr Arg Asn Leu Ser Ala Ala Ala Val Arg
435 440 445

Gly Glu Tyr Glu Thr Cys Val Leu Tyr Ser Asp Gln Leu Lys Gln Arg
450 455 460

Cys Val Asp Gln Phe Lys Glu Gly Ser Leu Gly Ile Glu His Leu Ala
465 470 475 480

Ala Val Asp Ser Ile Cys Asp Phe Val Ser Lys Ala Met Gly Ala Ala
485 490 495

Asp Pro Ile Arg Thr Tyr His Val Asn Leu Ser Ile Phe Thr Ser Ile
500 505 510

Pro Asp Phe Trp Ala Phe Gly Gln Leu Phe Pro Ile Val Pro Ile His
515 520 525

Arg Leu Asp Glu Lys Pro Ala Val Arg Gly Ile Leu Ser Asp Leu Thr
530 535 540

Cys Asp Ser Asp Gly Lys Val Asp Lys Phe Ile Gly Gly Glu Ser Ser
545 550 555 560

Leu Gln Leu His Glu Leu Gly Ser Asn Gly Asp Gly Gly Tyr Tyr
565 570 575

Leu Gly Met Phe Leu Gly Gly Ala Tyr Glu Glu Ala Leu Gly Gly Leu
580 585 590

His Asn Leu Phe Gly Gly Pro Ser Val Val Arg Val Val Gln Ser Asp
595 600 605

Ser Ala His Ser Phe Ala Met Ser Arg Ser Val Pro Gly Pro Ser Cys
610 615 620

Ala Asp Val Leu Arg Ala Met Gln His Glu Pro Glu Leu Met Phe Glu
625 630 635 640

Thr Leu Lys His Arg Ala Glu Glu Phe Leu Glu Gln Glu Glu Asp Lys
645 650 655

Gly Leu Ala Ile Ala Ser Leu Ala Ser Ser Leu Ala Gln Ser Phe His
660 665 670

Asn Met Pro Tyr Leu Val Ala Pro Ala Ser Cys Cys Phe Thr Ala Val
675 680 685

Thr Ala Asn Asn Gly Gly Tyr Asn Tyr Tyr Tyr Ser Asp Glu Asn Ala
690 695 700

Ala Asp Ser Ala Thr Gly Glu Asp Glu Ile Trp Ser Tyr Cys Thr Ala
705 710 715 720

<210> 25

<211> 914

<212> DNA

<213> Plant

<400> 25

aagctcgann ttaancctca ntaaaaggaa caaaaagctgg taccgnggcc cccccctcgag 60
gtcgacggta tcgataaagct tgattaagct tagtangcac attagcagcg cttggatga 120
ttttaggcgc ggcctattcc ctggctat ataatcggt gggtctggga attaaaaccc 180
gatttcctcc ataaaattctc cgatctaaat ggcanagaat ttcatatattt atacctttc 240
ttgttgagtttggatg ggtgtttacc ccaaagtgtt cctggactg catgcataca 300
tccgtaagta acttaagtgc aacatggaaa atttcattga gaggaatcg caaaaaaaaaa 360
aagcttaatc gaattcctgc agccccgggg atccactagt tctagagcgg ccggccaccgc 420
ggtggagctc caattcgcgg tatagtgagt cgtattacaa ttcaactggc cgtcggttta 480
caacgtcgtg actggggaaan gcctggcggtt accaacttaa tcgccttgca gcacatcccc 540
ctttcgccag ctgggcgtaa tagcgaanag gcccgcacgt cgccctccca acagttgcgc 600
acctgatggn gaatgggacn gccctgtanc nccgcangaa ncgcggcngg tgggtggta 660
ccccancgtg acgcnaactt gcaacgccta acgcgcncct tcgccttc 720
aagttcnccgg ctteccgtaa gtccaaagcgg gggnccttag gttcgattat gttnggaccc 780
ccccnaaact gttanggtnt gtgatttggc accccaaaaac gttcccttg nctggtcntt 840
ttaaangcct ntcaacngaa accaccatcg gnacttgtt aagntncctt gcntgnaaaa 900
nngtaaattttnnn 914

<210> 26

<211> 829

<212> DNA

<213> Plant

<400> 26

agcaagctcg atatcgccct cactaaagg aacaaaaact ggtaccgggc cccccctcgta 60
ggtcgacggat atcgataaagc ttgattaanc tttttttttt tgaactacac aaggaaattt 120
cttctcctna gtaacacatg agaataatta gtgcaataaa ttacaagagg aacattgcag 180
ttggatttaaa gaatctgcgc tgggaattt agcctcaata ttgcataaa ccgtacagat 240
ttcaactgcatt tcatgaacga tagtatccgt gacacatcct tttggatgcc gtcctgtcca 300
catatgccac tactcacatc cactccattt ggttaagtt gcagaaagag cttcacaaac 360
attctccggg ttaattcctc ctgccaagag ccaccatgt ttgcctctaa tgccggcag 420
cttaaactga acccagttga atcccttgcc actgccacct ttgcactat ccactagaac 480
ccaatcaacc agagaagact cctcatcaga aatatagttc aaaaggctcc tcttcatttg 540
catgaagtac gtatattaca cgtttttccc tgaccaaagc ttaatcgaaat tctgcagccc 600
gggggatcng gnattctaga gcccgcac ggggtggagc tccaatcgcc taaatgancc 660
ataaaaatcac tggccgtcg ttanacgncc ggacggaaa cctgggtacc aacttaatcg 720
cctgnagcna tccccctcnc agcggngtan acgaaaggcc gncgattgcc tccanattgc 780
cacnggatgg aanggacncc gtnccganga acnggggnnn ggggtaccc 829

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/12450

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) A01H 5/00; C07H 21/04; C12N 5/14, 15/29, 15/52, 15/82
US CL 435/320.1, 414, 419; 536/23.2, 23.6, 24.5; 800/278, 317.3

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/320.1, 414, 419; 536/23.2, 23.6, 24.5; 800/278, 317.3

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HASHIMOTO et al. Intraspecific Variability of the Tandem Repeats in Nicotiana Putrescine N-methyltransferase. Plant Molecular Biology. 1998, Vol. 37, pages 25-37, especially Figure 3.	12 ----
---		15,16
Y		
X	HIBI et al. Gene Expression in Tobacco Low-Nicotine Mutants. The Plant Cell. May 1994, Vol. 6, pages 723-735, especially Figure 3.	12 ----
---		15,16
Y		
X	IZHAKI et al. A Petunia cDNA Encoding S-Adenosylmethionine Synthetase. Plant Physiology. 1995, Vol. 108, pages 841-842, see entire article.	12 ----
---		15,16
Y		

<input checked="" type="checkbox"/>	Further documents are listed in the continuation of Box C.	<input type="checkbox"/> See patent family annex.
* Special categories of cited documents:		
"A" document defining the general state of the art which is not considered to be of particular relevance	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"B" earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"&"	document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
17 AUGUST 2000	04 OCT 2000

Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer AMY NELSON Telephone No. (703) 308-0196
-------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/12450

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LAMATTINA et al. RNA Editing of the Transcript Coding for Subunit 4 of NADH Dehydrogenase in Wheat Mitochondria: Uneven Distribution of the Editing Sites Among the Four Exons. Nucleic Acids Research 1991, Vol. 19, No. 12, pages 3275-3282, especially Figure 4.	12
---		----
Y		15,16
X	LI et al. Arabidopsis Phosphoribosylanthranilate Isomerase: Molecular Genetic Analysis of Triplicate Tryptophan Pathway Genes. The Plant Cell. April 1995, Vol. 7, pages 447-461, especially Figure 3, page 459.	12,15
---		----
Y		16

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/12450

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
1-15,18-20
4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

The additional search fees were accompanied by the applicant's protest.
 No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/12450

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

STN, AGRICOLA, CAPLUS, BIOSIS, EMBASE, USPAT
search terms: putrescine methyltransferase, adenosylmethionine synthetase, ornithine decarboxylase, arginine decarboxylase, NADH dehydrogenase, phosphoribosylanthranilate isomerase, DNA, cDNA, gene, nucleic

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-16, drawn to coding DNA, vector, host cell, transgenic plant.

Group II, claim(s) 17, drawn to protein.

Group III, claim(s) 18-20, drawn to transformation method and transgenic plant with promoter DNA.

The inventions listed as Groups I, II, and III do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The coding DNA of Group I, e.g. Claim 12, is disclosed in the prior art publication of Hashimoto *et al.* (Plant Mol. Biol. 37: 25-37, 1998; see Fig. 3b). Therefore, there is no special technical feature which links the coding DNA of Group I with the protein of Group II.

Furthermore, there is no special technical feature under PCT Rule 13.2 which links the coding DNA of Group I and the transformation method and transgenic plant with the promoter DNA of Group III. Therefore, the inventions of Groups I, II, and III do not relate to a single inventive concept under PCT Rule 13.1.